A0610-1A

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

EDITORIAL	
ELEANOR SINGER, NANCY A. MATHIOWETZ, and MICK P. COUPER	Privacy, Confidentiality, and the 1990 U.S. Census
DIANA C. MUTZ	Routes to Politicizing Personal Experience
DONALD PHILIP GREEN and ERIC SCHICKLER	Multiple Measures of Party ID
YAACOV SCHUL and MIRIAM SCHIFF	Measuring Satisfaction with Organizations
COLM A. O'MUIRCHEARTAIGH, GEORGE D. GASKELL, and DANIEL B. WRIGHT	The Impact of Intensifiers
PAMI DUA and DAVID J. SMYTH	Public Pessimism about Unemployment
THE POLLS	Review: Rush to Judgment? *by Dianne Rucinski* Trends: Trends in Media Usage *by William G. Mayer*
BOOK REVIEWS	Reviews by *Stanley Feldman, John P. Robinson, Richard A. Brody, James A. Davis,* and *Carol A. Cassel*

Winter 1993

BEST AVAILABLE COPY

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Eleanor Singer, Chair (1991–93)
Gladys Engel Lang (1989–93)   W. Russell Neuman (1990–94)
Irving Crespi (1991–95)   J. Ronald Milavsky (1992–96)

Warren Mitofsky and Eleanor Singer, AAPOR Councillors-at-Large

---

Howard Schuman, *Editor*

Office Manager
Carlotta Nosel

Assistant to the Editor
Katherine Bischoping

The Polls
Reviews: Peter V. Miller, Albert E. Gollin
Trends: Robert Y. Shapiro

EDITORIAL BOARD

Herbert I. Abelson
Larry M. Bartels
Nathaniel Beck
Sandra H. Berry
Charles J. Brody
Pamela Conover
Jean M. Converse
Myra Marx Ferree
Steven E. Finkel
Richard J. Fox
Doris Graber
Graham Kalton
John P. Katosh
Gary King
Richard A. Kulka
Jacob Ludwig
William J. McGuire
Elizabeth Martin
Peter V. Miller
Diana C. Mutz
Helmut Norpoth
Vincent Price
Willard L. Rodgers
Nora Cate Schaeffer
Jacqueline Scott
Lee Sigelman
Tom W. Smith
Charlotte Steeh
John Sullivan
Marylee C. Taylor
Roger Tourangeau
Michael W. Traugott
Michael Weeks
Susan Welch
John Zaller

Public Opinion Quarterly (ISSN 0033-362X) is published quarterly in Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Ave., Chicago, IL 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $45.00; individuals 1 year $22.00. Special rates available to members of AAPOR. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $19.00. In Canada, add 7% GST to subscription price. Outside of U.S.A., add $3.00 for each year's subscription to cover postage. Japanese subscription agent: Kinokuniya Company Ltd. Individuals have the option to order directly from the University of Chicago Press.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to Journals Fulfillment Department, University of Chicago Press, P.O. Box 37005, Chicago, IL 60637. FAX: (312) 753-0811.

Single copy and back volume information furnished upon written request. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Institute for Social Research
P.O. Box 1248, Ann Arbor, MI 48106
Business, Subscriptions, Back Issues: University of Chicago Press,
P.O. Box 37005, Chicago, IL 60637
AAPOR: P.O. Box 1248, Ann Arbor, MI 48106

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

© 1993 by the American Association for Public Opinion Research

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ∞

PUBLIC OPINION QUARTERLY

VOLUME 57 / WINTER 1993 / NUMBER 4

Editorial iii

Articles

The Impact of Privacy and Confidentiality Concerns on Survey Participation: The Case of the 1990 U.S. Census 465
ELEANOR SINGER, NANCY A. MATHIOWETZ, and MICK P. COUPER

Direct and Indirect Routes to Politicizing Personal Experience: Does Knowledge Make a Difference? 483
DIANA C. MUTZ

Multiple-Measure Assessment of Party Identification 503
DONALD PHILIP GREEN and ERIC SCHICKLER

Measuring Satisfaction with Organizations: Predictions from Information Accessibility 536
YAACOV SCHUL and MIRIAM SCHIFF

Intensifiers in Behavioral Frequency Questions 552
COLM A. O'MUIRCHEARTAIGH, GEORGE D. GASKELL, and DANIEL B. WRIGHT

Survey Evidence on Excessive Public Pessimism about the Future Behavior of Unemployment 566
PAMI DUA and DAVID J. SMYTH

The Polls

Review: Rush to Judgment? Fast Reaction Polls in the Anita Hill–Clarence Thomas Controversy 575
DIANNE RUCINSKI

Trends: Trends in Media Usage 593
WILLIAM G. MAYER

Book Reviews

Benjamin I. Page and Robert Y. Shapiro, *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences* 612
STANLEY FELDMAN

Vincent Price, *Public Opinion* 614
JOHN P. ROBINSON

Samuel L. Popkin, *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns* 617
RICHARD A. BRODY

William G. Mayer, *The Changing American Mind: How and Why American Public Opinion Changed between 1960 and 1988* 619
JAMES A. DAVIS

Ruy A. Teixeira, *The Disappearing American Voter* 621
CAROL A. CASSEL

Index to Volume 57, 1993 623

Notice to Contributors

Prospective contributors should refer to the detailed Notice to Contributors that appears in the Spring 1993 issue of *POQ*.

THE IMPACT OF PRIVACY AND CONFIDENTIALITY CONCERNS ON SURVEY PARTICIPATION

THE CASE OF THE 1990 U.S. CENSUS

ELEANOR SINGER
NANCY A. MATHIOWETZ
MICK P. COUPER

Abstract This study investigates the relationship between concerns about confidentiality and privacy, on the one hand, and mail returns to the 1990 census, on the other. Such concerns significantly affect mail returns, even when demographic variables known to be related to both concerns and survey participation are controlled. However, their impact is not very large, and the effects of confidentiality and privacy concerns vary for black and white respondents.

Introduction

It is an article of faith among survey researchers that people will not respond to surveys, or at any rate not respond honestly, unless they are assured of the confidentiality of their replies (see, e.g., U.S. Department of Health, Education, and Welfare 1973, p. 93; cited in Dalenius 1983, p. 88). The evidence for this belief is at best modest (for a review of the evidence, see Singer 1993; Singer, Hippler, and Schwarz 1992). While some early field experiments suggest modest effects on either survey response rates or response quality, or both, we know little about the mechanisms by which these effects come about. In

ELEANOR SINGER is a senior social scientist at Columbia University and a social science analyst at the U.S. Bureau of the Census. NANCY A. MATHIOWETZ is deputy director of the Division of Statistics and Research Methodology at the Agency for Health Care Policy and Research. MICK P. COUPER is a research investigator at the Survey Research Center, University of Michigan, and visiting researcher at the U.S. Bureau of the Census. The views expressed in this article are those of the authors, and no official endorsement by the U.S. Bureau of the Census or the Agency for Health Care Policy and Research is intended or should be inferred.

Public Opinion Quarterly Volume 57:465–482 © 1993 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/93/5704-0004$02.50

fact, recent experiments by Singer, Hippler, and Schwarz (1992) suggest that under some circumstances, an elaborate assurance of confidentiality may be counterproductive, increasing respondents' concerns and their reluctance to respond.

There is also a large and growing literature on public concerns about privacy (see, e.g., Flaherty 1989; Louis Harris and Associates 1983; Marx 1988; Westin 1967, 1981, 1990). Unlike concerns about data confidentiality, which can in principle be allayed by an appropriate assurance, concerns about privacy go to the heart of a respondent's willingness to participate in a survey. By concern about confidentiality, we refer to a desire to keep data already given to one agent out of the hands of others; by concern about privacy we refer to a desire to keep information about oneself out of the hands of others altogether. So far as we know, no prior studies have investigated the role of privacy concerns as factors in actual, rather than self-reported, survey response.

What earlier studies of the role of confidentiality in survey response have in common is manipulation of an introduction containing an assurance of confidentiality as the independent variable. With the exception of a few recent studies, they have not investigated the underlying assumptions—namely, that concern about the confidentiality of the data reduces the likelihood of response, and that an assurance of confidentiality reduces this concern.

The present study seeks to remedy the lack of research on these assumptions, at least in part. Its purpose is to investigate the relationship between concerns about confidentiality and privacy, on the one hand, and mail return to the 1990 census, on the other. We do not, however, vary the kind of confidentiality assurance given to respondents and so can say nothing about whether it is possible to reduce the concerns expressed. Unlike previous investigators (Fay, Bates, and Moore 1991; Fay, Carter, and Dowd 1991; and Kulka et al. 1991), we use actual returns, rather than self-reported returns, as the dependent variable. Since returning a census form, like voting in an election, may indicate good citizenship, self-reports may be subject to social desirability effects. The use of actual mail returns eliminates this possibility.

Prior Research on Confidentiality and Privacy Concerns and the 1990 Census

The decennial census is in no way a typical survey; its mandatory response, the ability to utilize media to enhance response, and its use for determining congressional districts and the allocation of federal funds all suggest that response to the census should be buffered against the trend toward higher nonresponse evident in most surveys during the 1980s. There is evidence, however, that in 1990 the decline in mail response to the decennial census followed the same trend that has been evident in academic and marketing surveys in the United States over the past decade, and even longer (Groves 1989; Steeh 1981). Plans for the 1990 census had assumed a 70 percent mail response rate; the comparable response rate in 1980 was approximately 75 percent. By the start of nonresponse field follow-up, however, the 1990 mail response rate had reached only 64.6 percent.[1] A multitude of hypotheses concerning possible explanations for the low response has been suggested, ranging from the rise in illiteracy rates among the adult population to concerns about confidentiality and privacy.

Two studies carried out in 1990—the Outreach Evaluation Survey, conducted by the Census Bureau, and the Survey of Census Participation, carried out by the National Opinion Research Center (NORC)—have examined the role of privacy or confidentiality concerns in self-reported response to the 1990 census. Fay, Bates, and Moore (1991), for example, reported that the large majority of respondents to the Outreach Evaluation Survey "appear to believe that [census] data are kept confidential" (p. 18). Furthermore, they state that "this belief has not decreased since the last decennial census," making such a decline an unlikely explanation for the observed decline in the mail-back response rate.

But although the level of trust in the promise of confidentiality has apparently not declined, the relationship between trust and self-reported mail response has undergone some change between 1980 and 1990. Whereas trust in the Census Bureau's assurance of confidentiality, as measured by an index of three identical items across both years, was not predictive of self-reported census mail response in 1980, it was predictive of such response in 1990 (Fay, Bates, and Moore 1991, p. 18 and table 4). Of those with low confidence in the Census Bureau in 1980, 90.9 percent said they had mailed back their census form, compared with 90.7 percent of those with high confidence. In 1990, however, only 73.5 percent of those with low confidence, compared with 86.1 percent of those with high confidence, reported having mailed back their census form.

---

[1] The Census Bureau uses two response rate calculations. The "response rate," as defined by the Census Bureau, indicates the percentage of forms that are completed and mailed back as a percentage of all the forms distributed, thus including vacant housing in the denominator. The complementary nonresponse rate therefore defines the level of field follow-up necessary, since vacant housing must be visited by an enumerator to verify the vacant status. The "return rate," as defined by the Census Bureau, more closely parallels response rates reported by most surveys, in which the numerator indicates the number of completed and returned forms and the denominator indicates the number of occupied housing units eligible to return forms. In 1990, the census return rate was 74.1 percent.

Thus, while the public's confidence in the confidentiality of data maintained by the Census Bureau may not have declined between 1980 and 1990, those who entertained doubts were more likely to withhold cooperation in the latter year. Kulka et al. (1991), examining predictors of self-reported mail returns in the Survey of Census Participation, also identified concerns about privacy and confidentiality as factors in self-reported participation.

Although the Outreach Evaluation Survey turned up no evidence of declining confidence in the Census Bureau's assurance of confidentiality, surveys conducted outside the Census Bureau have documented a substantial increase in concern over privacy (as distinct from confidentiality) over time. Harris, for example, reported that 31 percent of a national sample indicated they were "very concerned" about privacy in 1978; by 1983, that percentage had increased to 48 percent and remained at 46 percent as of 1990 (Westin 1990). Westin (1990, table 1-1) also reported a number of demographic correlates of concern about privacy: females, black respondents, those with less education, and those with lower income were particularly likely to express such concern.

In this article, we extend the investigation of privacy and, especially, confidentiality as factors in survey response by examining their role in actual mail return to the 1990 census. We also examine demographic predictors of concern about privacy and confidentiality and control for these in the analysis.

Methodology

SURVEY DESIGN

As a result of the low mail response rate to the 1990 decennial census, the Bureau of the Census contracted with NORC to conduct a nationwide survey focusing on issues related to census participation. The questionnaire for the Survey of Census Participation (SCP) was structured to collect information about a number of hypotheses that had been cited in focus group research as reasons for not responding to the census—for example, the effects of the format of the census form, concerns with privacy and the confidentiality of the data, negative attitudes toward monolithic government, and lack of knowledge concerning the uses of census data. The questionnaire also included questions concerning general receipt and handling of mail within the household, involvement in community activities, prior experience with surveys, attitudes toward answering survey questions, and demographic characteristics of respondents and other household members.

The sample was a national probability sample of 2,760 households. Data collection took place between June and August 1990. Face-to-face interviews were conducted with the person in the household identified as having the most to do with the census form or the mail in general. Information was collected from 2,478 respondents for an overall response rate of 89.8 percent.

Data from the SCP were linked to decennial census information, thus permitting an analysis of the role of privacy and confidentiality concerns (as reported in the SCP) in census participation. Among the 2,478 SCP respondents, an exact match was made for 2,418 households, or 97.6 percent. Matches were completed at the address and not the person level. For a case to be considered a match, correspondence on all aspects of the two addresses (city/town, street name and number, unit or apartment number, description, etc.) was required.[2] Because matches were made to current addresses, whole households that had moved between the time of the census and the SCP data collection were not included in the analyses (a total of 149 SCP respondents).

For purposes of the present analyses, response to the decennial census is defined as those individuals who were targeted for mail-out, mail-back procedures and who completed the census questionnaire and returned the form without a personal visit from a census enumerator. Approximately 95 percent of the U.S. population was targeted for mail-out, mail-back procedures. Those in the SCP sample not targeted for such procedures were excluded from the analyses ($N = 115$); those who failed to return their form by mail are classified as nonrespondents, even though information for these households may have been obtained eventually through personal visit and other methods of enumeration. These exclusions (nonmatches, whole-household movers, and those not targeted for enumeration by mail) leave us with 2,183 eligible sample households for analysis. The estimated return rate for the SCP respondents used in this analysis is 76.1 percent, very close to the 74.1 percent return rate estimated for the U.S. population as a whole from the decennial files.

One of the major difficulties in studying nonresponse, in this case to the decennial census, is the lack of attitudinal information for nonrespondents. In the SCP data collection, more than 10 percent of the sampled cases were classified as nonrespondents, either because of refusals or because of noncontact with a household member. To the extent that the relationship between concerns about privacy and confidentiality and census participation differs between SCP respondents and nonrespondents, estimates of the effects of these concerns on census participation may be biased. However, since the nonresponse rate was low, we believe that this bias is likely to be small.[3]

---

[2] The match of survey and census data was done as part of a larger project on survey participation (see Groves and Couper 1992). For details of the match operation, see Couper and Groves 1992.

As noted earlier, the dependent variable of interest, mail return to the 1990 census, is based on actual rather than self-reported information. A comparison of actual versus self-reported response indicates that a small percentage of SCP respondents misreported response to the census. Among SCP respondents included in our analysis, 8.8 percent overreported that they had completed and returned the form by mail, and an additional 2.1 percent stated that they had not completed the form when, according to decennial records, the census form had been returned. This yields a net overreport of 6.7 percent. Investigation of the extent to which the relationships we report differ for self-reported versus actual behavior is beyond the scope of this article.

All results reported here are based on analyses of weighted data, reflecting unequal probabilities of selection. Tests of statistical significance and standard error estimates are based on Taylor Series approximation, reflecting the complexity of the design, using SUDAAN (Shah et al. 1991).

MEASURES

As noted above, the SCP contained several measures related to confidentiality and privacy. These measures included both items related to general concerns and items quite specific to the decennial census. For the sake of parsimony, exploratory factor analyses were completed to determine the critical variables for creating measures of privacy and confidentiality and to confirm the hypothesis that privacy and confidentiality indeed represent two different concepts.

Over thirty variables were included in an orthogonal factor analysis with Varimax rotation. Four factors resulted, two of them clearly defining privacy and confidentiality concerns.[4] On the basis of the factor analysis, we constructed a privacy index consisting of eight items and a confidentiality index consisting of six items, giving each item equal weight. The items making up each index are shown in Appendix A. The average factor loading score for the items making up the privacy index was .49, with a range between .27 and .62. For the items making up the index of confidentiality, the average factor loading score was .42, with a range from .27 to .65. Cronbach's alpha coefficient was calculated for each index as a measure of reliability. For the 8-item privacy index, alpha is .726: for the index of confidentiality, alpha is .559. Although the results of the factor analysis support the conceptual distinction between privacy and confidentiality, there is a moderate (.251) intercorrelation between the two indexes. It should be noted, as well, that the items making up the confidentiality index all relate specifically to trust in the Census Bureau's handling of the data, whereas all but two of the questions on the privacy index relate to more general concerns about privacy.

---

[3] The study by Groves and Couper (1992) includes an investigation of nonresponse to the Survey of Census Participation.

[4] Using Rasch modeling techniques to analyze knowledge and attitude items on the SCP, Fay, Carter, and Dowd (1991) demonstrate that "violation of privacy" is not closely linked to confidentiality concerns but conclude that "there is not a single dimension corresponding to respondents who are more or less knowledgeable about confidentiality" (p. 530). Although the Varimax rotation does produce a single factor, the reliability of the resulting index, as measured by alpha, is at best moderate.

Table 1 shows the distribution for the two indexes. A low score on each index indicates little or no concern and a high score indicates intense concern.

Results: Privacy and Confidentiality as Predictors of Census Mail Return

The bivariate relationships between confidentiality concerns and census mail return, and privacy concerns and census mail return, are

Table 1. Distributions for Privacy and Confidentiality Indexes

Index Scores	Privacy Index (%)	Confidentiality Index (%)
0 (Low concern)	23.1	23.2
1	22.1	30.7
2	19.5	22.8
3	12.4	13.7
4	9.3	5.9
5	6.0	2.6
6	4.1	1.1
7	2.1	. . .
8 (High concern)	1.4	. . .
Total	100.0	100.0

shown in table 2. Because of the skewness of the distributions, the two highest categories on both indexes were collapsed. Scores on both indexes are significantly related to census mail return. Neither relationship, however, is a simple linear one. Instead, the decline in response comes only among those whose scores fall at or above the midpoint of the respective indexes.

Linear regression models of demographic characteristics on privacy and confidentiality indexes are shown in table 3. The demographic variables included in the table are those that, on the basis of theoretical considerations or prior empirical work, have a plausible connection to concerns about either privacy or confidentiality. For example, Westin (1990, table 1-1) found that women evinced greater privacy concerns than men; those with less education were more concerned about privacy than those with more education; and blacks expressed more concern about privacy than Hispanics or whites. We also reasoned that those whose primary language was other than English or who were not born in the United States might be more concerned about data confidentiality because of concerns about immigration laws, and that those living in households with unrelated individuals might be especially sensitive to the privacy of those individuals. Those with children in the household might be especially concerned about privacy because

Table 2. Census Mail Return Rates by Concerns with Privacy and Confidentiality

	Percentage Returning Census Form	
Index Scores	Privacy Index	Confidentiality Index
0	77.6	77.5
1	78.7	78.6
2	78.5	78.9
3	75.9	70.1
4	74.2	72.7
5[a]	69.4	54.8
6	66.8	
7[b]	63.2	
χ^2	16.9	15.8
df	7	5
p	<.05	<.01

[a] Scores of 5–6 were collapsed on the confidentiality index.
[b] Scores of 7–8 were collapsed on the privacy index.

Table 3. Linear Regression Coefficients of Demographic Characteristics on Privacy and Confidentiality Indexes

	Privacy		Confidentiality	
Predictors[a]	Coefficient	Standard Error	Coefficient	Standard Error
Intercept	1.975**	.374	1.758**	.256
Age	−.008**	.003	.0002	.002
Gender (male)	.120	.084	.218**	.068
Education	−.123**	.046	−.184**	.033
Language (English)	.655**	.221	.230	.223
U.S. born	−.163	.156	.023	.128
Race/ethnicity:				
Hispanic	.116	.156	−.047	.149
Black non-Hispanic	.779**	.169	.313*	.122
White non-Hispanic	. . .	. . .	. . .	. . .
Household characteristics:				
Household size	.079	.048	.016	.025
Children in household	−.421**	.135	−.046	.093
All related	.148	.133	−.026	.087
SMSA size:				
City	.074	.116	−.105	.098
Suburb	.048	.197	.111	.137
Rural	. . .	. . .	. . .	. . .
R^2	.031		.036	
N	2,100		2,100	

[a] See App. B for variable descriptions.
* $p < .05$.
** $p < .01$.

of a desire to protect the children's safety; at the same time, the presence of children was likely to reduce a household's privacy (and, therefore, perhaps reduce privacy concerns as well) because children serve to link the household to a variety of community organizations: schools, churches, youth groups, recreational organizations, and the like (cf. Groves and Couper 1993). A variety of theoretical considerations (see, e.g., Simmel 1955) led us to expect greater concerns about privacy among those living in urban as opposed to other areas.

The models in table 3 include demographic characteristics of the SCP respondent (and therefore the respondent to the census), characteristics of the household, and two dummy variables reflecting standard metropolitan statistical area (SMSA) size, with the rural classification as the omitted category. (See App. B for a complete description of the independent variables.) As can be seen from table 3, education and race are significantly related to both indexes, whereas age, English-speaking status, and the presence of children in the household are related to privacy concerns only, and gender is related to confidentiality concerns only. Those without children in the household, and younger respondents, were more concerned about privacy; men were more likely to voice concerns about confidentiality than women. Surprisingly, respondents whose primary language was English were significantly more likely to voice concerns with respect to privacy. Like Westin (1990), we see that black non-Hispanic respondents are more concerned with privacy than either Hispanic or white non-Hispanic respondents, and black non-Hispanic respondents exhibited higher levels of concern about confidentiality as well. Those with more education were less likely to be concerned about either issue. Although women in Westin's sample expressed more concern about privacy than men, in our data there is a nonsignificant tendency in the other direction.

To safeguard against the possibility that the relationships in table 2 between privacy and confidentiality concerns and census mail return rates are spurious, due to the effect of one or more demographic characteristics on both privacy and confidentiality concerns and census participation, we ran a series of nested logistic regression models in which census mail return is predicted as a function of privacy and confidentiality, controlling for demographic characteristics. The models are shown in table 4. The first model predicts census mail return from a set of demographic characteristics only; the second predicts mail return from both privacy and confidentiality indexes in addition to the set of demographic characteristics.

As can be seen from the first column of table 4, age, gender, education, race, and whether all members of the household are related all significantly predict census mail return. Those who are older, better educated, and live in households all of whose members are related are significantly more likely to mail back their census form; men are significantly less likely to do so than women, and black non-Hispanic respondents significantly less likely to do so than white non-Hispanic or Hispanic respondents.

A test of the joint significance of confidentiality and privacy reveals that the two variables together make a significant ($p = .001$) contribution to the explanation of the dependent variable. With both variables in the model, the confidentiality coefficient appears slightly stronger than the privacy coefficient (relative to their respective standard errors). This finding accords with our impression from the bivariate relationships (cf. table 2) that specific distrust in the confidentiality of data collected by the Census Bureau is a stronger predictor of census mail return than mere diffuse concerns about privacy violations.

Table 4. Logistic Regression Coefficients of Demographic Characteristics and Privacy and Confidentiality Concerns on Census Mail Return

	Demographics Only		Demographics, Privacy, and Confidentiality	
Predictors[a]	Coefficient	Standard Error	Coefficient	Standard Error
Intercept	−2.469**	.457	−2.187**	.474
Age	.040**	.004	.040**	.004
Gender (male)	−.348**	.123	−.322**	.123
Education	.343**	.057	.319**	.058
Language (English)	.447	.446	.525	.457
U.S. born	.467	.251	.460	.252
Race/ethnicity:				
Hispanic	−.018	.255	−.013	.266
Black non-Hispanic	−.918**	.165	−.844**	.168
White non-Hispanic	. . .	. . .	. . .	. . .
Household characteristics:				
Household size	.017	.046	.024	.046
Children in household	.044	.199	.013	.201
All related	.577**	.132	.589**	.133
SMSA size:				
City	.040	.152	.028	.150
Suburb	−.190	.206	−.177	.207
Rural	. . .	. . .	. . .	. . .
Privacy index			−.060*	.030
Confidentiality index			−.112*	.044
N	2,100		2,100	

[a] See App. B.
* $p < .05$.
** $p < .01$.

Because of our concern that the relationship between privacy and confidentiality, on the one hand, and census return, on the other, was not a simple linear one, we fitted a number of alternative transformations of the privacy and confidentiality indexes in addition to those shown in table 4. These alternative transformations included using the natural log of confidentiality and fitting a quadratic transformation of both the confidentiality and the privacy indexes. None of these transformations significantly improved the fit when compared with the second model in table 4.

However, because black respondents exhibited greater privacy and confidentiality concerns as well as lower census mail returns, we estimated separate logistic regression models for black and white respondents. Because of their relatively small sample size, Hispanics were excluded from the analysis. Table 5 presents the results of these separate models.

In contrast to the interpretation of the second model in table 4, which would indicate that there is a significant effect of privacy and confidentiality for both blacks and whites, the first two columns of table 5 suggest that these effects are significant only for white respondents, among whom higher concerns about both privacy and confidentiality lead to lower census returns. Although black respondents express higher levels of concern about privacy and confidentiality than white respondents, these higher levels of concern are not the cause of their lower rate of mail return. Controlling for other variables in the model, the effects of privacy and confidentiality are insignificant, and the direction of the effect associated with confidentiality suggests that higher levels of concern actually result in higher census return rates.

Examination of the bivariate relationship between confidentiality concerns and census mail return rates for the black population indicated that the relationship appeared to be curvilinear. Return rates for either very low (scores of zero) or very high levels of concern on the confidentiality index averaged around 45 percent; among blacks with confidentiality scores of two or three, census return rates averaged 68 percent. Column 3 of table 5 presents the results of fitting the quadratic form of confidentiality for blacks to test the curvilinearity hypothesis. The addition of the quadratic transformation significantly improves the fit ($p < .01$) as compared to the linear model for blacks. A similar examination of the relationship between scores on the privacy index and census return indicates that there is essentially no effect of privacy concerns for black respondents.

Similar tests were performed on the whites-only model, but the results support the linearity assumption for this group. Neither the squared term for confidentiality nor the squared term for privacy improved the fit of the model for whites.

In addition to the complex effects of confidentiality on census mail return for blacks, we note that the relationship between several person-level and household-level characteristics and census return changes (from the direction and significance presented in table 4) when separate models are estimated for the black and white populations. Gender has no significant effect among black respondents, and whether everyone in the household is related also is insignificant. Whereas language had no effect in the overall model (table 4), there is a positive effect among whites and no effect among blacks. Household size (originally insignificant) has an insignificant negative effect on census return among

Table 5. Logistic Regression Coefficients of Demographic Characteristics and Privacy and Confidentiality Concerns on Census Mail Return Separately for Whites and Blacks

Predictors	White Respondents	Black Respondents: Linear Model	Black Respondents: Quadratic Model
Intercept	−2.799** (.634)	−5.286** (1.371)	−5.681** (1.335)
Age	.043** (.005)	.049** (.009)	.051** (.010)
Gender (male)	−.360** (.134)	−.010 (.317)	−.094 (.306)
Education	.347** (.077)	.480** (.142)	.506** (.134)
Language (English)	1.530* (.602)	−.732 (1.110)	−.966 (1.113)
U.S. born	.328 (.433)	.281 (.924)	.223 (.968)
Household characteristics:			
Household size	−.107 (.071)	.489** (.134)	.499** (.140)
Children in household	.348 (.248)	−.472 (.476)	−.504 (.502)
All related	.741** (.163)	.452 (.345)	.335 (.353)
SMSA size:			
City	−.088 (.175)	1.368* (.542)	1.505** (.546)
Suburb	−.309 (.234)	1.323 (.765)	1.541* (.744)
Rural	. . .	. . .	. . .
Privacy	−.099* (.040)	−.012 (.056)	−.007 (.055)
Confidentiality	−.165** (.058)	.134 (.109)	.896** (.289)
Confidentiality squared	. . .	. . .	−.168** (.062)
N	1,518	295	295

Note.—Standard errors are presented in parentheses.

* $p < .05$.

** $p < .01$.

whites and a strong positive effect among blacks. Similarly, SMSA size has no effect among whites, but living in either a city or a suburban area, as compared to living in rural areas, has a strong positive effect on census mail return among blacks.

While the quadratic model provides a better fit to the data, it does not provide an easily understandable explanation of the relationship between confidentiality and census return rates among the black population. A variety of analyses were undertaken in an effort to clarify this relationship. The factor structure of the confidentiality items is similar for whites and blacks, but the relationship between some of the individual items and census mail return differs between the two race groups. The interpretation of this relationship thus remains ambiguous and deserves attention beyond the scope of this article. However, it is safe to conclude that confidentiality concerns do not appear to be linearly related to census returns among black respondents. Among white respondents, on the other hand, greater distrust of the Census Bureau's safeguarding of confidentiality is significantly associated with a lower rate of census mail return.

The effect of adding the privacy and confidentiality indexes is not large in absolute terms—for whites, they explain approximately 1.3 percent of the variance in mail return rates over and above that explained by the demographic characteristics alone.[5] Nevertheless, for white respondents the predicted effect of a reduction of 2 points on the confidentiality index is an increase in the likelihood of returning a census questionnaire of approximately 3–6 percentage points, depending on other characteristics of the respondent and the exact confidentiality score.

Summary and Conclusion

This research sought to determine the impact of concerns about confidentiality and privacy on survey participation, specifically, response

---

[5] For an estimate of how much variance in census mail return is accounted for by attitudes about privacy and confidentiality, over and above the basic demographic model, we have used Hosmer and Lemeshow's (1989, p. 148) R_L^2, which can be considered a rough analogue of the measure of variance explained obtained from an ordinary least-squares regression. This measure is calculated as follows:

$$R_L^2 = \frac{-2\text{Log } L_0 - (-2\text{Log } L_1)}{-2\text{Log } L_0}$$

where log L_0 is the log likelihood of the null or intercept-only model, and log L_1 is the log likelihood of the fitted model (see also DeMaris 1992, p. 53). Using this approach we obtain an R_L^2 of 0.140 for the demographics-only model, and an R_L^2 of 0.153 for the full model.

to the 1990 mail census. Using data from the Survey of Census Participation, we isolated two separate factors, one measuring generalized concerns about privacy, the other more specific trust in the Census Bureau's maintenance of data confidentiality.

In multivariate models including these two measures in addition to demographic characteristics, both indexes have a significant effect on census mail return. However, the effect of privacy and confidentiality concerns differs for white and black respondents. For whites, a decrease in concern about privacy and confidentiality issues is associated with an increase in the probability of mailing back the census form. But although black respondents express greater concern about both privacy and confidentiality than white respondents do, there is no relationship between privacy concerns and mail return rates for black respondents, and the effect of confidentiality concerns is curvilinear: an increase in concern appears to be related, at least up to a point, with an increased tendency to respond.

Even among whites, the effect of privacy and confidentiality concerns is not large. But given the cost of retrieving census information not returned by mail, even an increase of 3–6 percentage points in the likelihood of returning the form by mail is by no means trivial. And if the issue of data confidentiality became more salient to the American public, one might expect the effects on nonresponse to be much more substantial. Such a hypothesis is suggested by focus group discussions about confidentiality issues held under the auspices of the U.S. Census Bureau (Singer and Miller 1992).

None of the demographic variables we examined predict either privacy or confidentiality concerns very well—together they explain about 3 percent of the variance in both scores. Although several factors were significant in predicting privacy and confidentiality scores, most notably the race of the respondent, the lack of explained variance suggests that demographic characteristics are not the critical variables determining either privacy or confidentiality concerns.

More generally, this article provides support for the assumption that, at least for white respondents, concern about privacy and about the confidentiality accorded one's data contributes significantly to survey nonresponse. However, the effects are not large, and so far as confidentiality is concerned, the crucial variable appears to be trust in the integrity of the data-collection agency, not the nature of the assurance given to respondents. Past research has focused on variations in confidentiality assurances, rather than on the respondent's trust in the confidentiality of the data. Although the two things are undoubtedly related, there is not necessarily a one-to-one correspondence between them. Future research on the role of confidentiality ought to take both variables into account.

Appendix A

Survey Questions

PRIVACY INDEX

Q29g. The Census is an invasion of privacy. (Agree/Disagree)
Q34. Do you think the government bothers you too much with requests for information? (Yes/No)
Q39. Do you ever feel your privacy is being violated by:
  A. Banks and credit companies when they ask about finances? (Yes/No)
  B. Neighbors who gossip about you and your family? (Yes/No)
  C. The government when it collects tax returns? (Yes/No)
  D. The government when it takes the census? (Yes/No)
  E. Computers which store a lot of information about you? (Yes/No)
  F. The people who ask questions on public opinion surveys? (Yes/No)

CONFIDENTIALITY INDEX

Q15d. Do the police and the FBI use the census to keep track of troublemakers? (Yes/No)
Q15f. How about to locate illegal aliens? Is the census used for that? (Yes/No)
Q16. After the Census Bureau collects the information about you, can anyone outside the Census Bureau look at it? (Yes/No)
Q17. Do you feel that other government agencies could obtain census information about you if they really tried? (Yes/No)
Q29c. The Census Bureau's promise of confidentiality can be trusted. (Agree/Disagree)
Q29h. People's answers to the census cannot be used against them. (Agree/Disagree)

Note: Each item in the indexes was recoded such that a zero represented the response indicating no concern and a value of one indicated the response that expressed concern.

Appendix B

Description of Independent Variables

Age: Measured as a continuous variable, ranging from 17 to 96.
Gender: Dummy-coded variable, where 1 = male and 0 = female.
Education: Ordinal variable indicating level of education of the respondent. The levels for the variable are less than high school graduate, high school

graduate, some college or other education beyond high school, college graduate, and postgraduate degree.

Primary Language: Dummy-coded variable, where 1 = English and 0 = any other language.

Household Size: Continuous variable indicating size of household.

Children in Household: Dummy-coded variable, where 1 = children under the age of 17 present in household and 0 = no children under 17.

All Related in Household: Dummy coded variable, where 1 = all members of household are related by blood, marriage, or adoption and 0 = not all related.

Race/Ethnicity: Coded as three mutually exclusive variables based on answers to questions concerning race and Hispanic origin. "Hispanic" indicates a positive response to the Hispanic origin question regardless of race. "Black, non-Hispanic" refers to individuals who listed their race as black and indicated "no" with respect to Hispanic origin. "White, non-Hispanic" includes all other race categories for which the person indicated no to the Hispanic origin question. The omitted category for the model is "white, non-Hispanic."

U.S. Born: Dummy-coded variable where 1 indicates that the respondent was born in the United States; coded 0 for all others.

SMSA Size: A series of dummy-coded variables indicating metropolitan area ("City"), suburban area contiguous to a metropolitan area ("Suburb"), and all other areas ("Rural"). The omitted category in the model is "rural."

References

Couper, Mick P., and Robert M. Groves. 1992. "Survey-Census Match Procedures." Unpublished memorandum, Bureau of the Census, Washington, DC.

Dalenius, Tore. 1983. "Informed Consent or R.S.V.P." In *Incomplete Data in Sample Surveys*, ed. W. G. Madow et al., 3:95–106. New York: Academic Press.

DeMaris, Alfred. 1992. *Logit Modeling: Practical Applications*. Newbury Park, CA: Sage.

Fay, Robert L., Nancy Bates, and Jeffrey Moore. 1991. "Lower Mail Response in the 1990 Census: A Preliminary Interpretation." In *Proceedings of the Annual Research Conference*. Washington, DC: Bureau of the Census.

Fay, Robert L., Woody Carter, and Kathryn Dowd. 1991. "Multiple Causes of Nonresponse: Analysis of the Survey of 1990 Census Participation." In *Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association.

Flaherty, David H. 1989. *Protecting Privacy in Surveillance Societies*. Chapel Hill: University of North Carolina Press.

Groves, Robert M. 1989. *Survey Error and Survey Costs*. New York: Wiley.

Groves, Robert M., and Mick P. Couper. 1992. "Correlates of Nonresponse in Personal Visit Surveys." In *Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association.

———. 1993. "Unit Nonresponse in Demographic Surveys." In *Proceedings of the Annual Research Conference*. Washington, DC: Bureau of the Census.

Hosmer, David W., and Stanley Lemeshow. 1989. *Applied Logistic Regression*. New York: Wiley.

Kulka, Richard, Nicholas Holt, Woody Carter, and Kathryn L. Dowd. 1991. "Self Reports of Time Pressures, Concerns for Privacy and Participation in the 1990 Mail

Census." In *Proceedings of the Annual Research Conference*. Washington, DC: Bureau of the Census.

Louis Harris and Associates. 1983. *The Road after 1984*. New York: Louis Harris.

Marx, Gary T. 1988. *Undercover: Police Surveillance in America*. Berkeley and Los Angeles: University of California Press.

Shah, B. V., B. G. Barnwell, P. N. Hunt, and L. M. LaVange. 1991. *SUDAAN User's Manual, Release 6.0*. Research Triangle Park, NC: Research Triangle Institute.

Simmel, Georg. 1955. *Conflict and the Web of Group Affiliation*. New York: Columbia University Press.

Singer, Eleanor. 1993. "Informed Consent and Survey Response: A Summary of the Empirical Literature." *Journal of Official Statistics* (in press).

Singer, Eleanor, Hans-Juergen Hippler, and Norbert Schwarz. 1992. "Confidentiality Assurances in Surveys: Reassurance or Threat?" *International Journal of Public Opinion Research* 4:256–68.

Singer, Eleanor, and Esther R. Miller. 1992. "Reactions to the Use of Administrative Records: Results of Focus Group Discussions." Bureau of the Census, Center for Survey Methods Research, Washington, DC.

Steeh, Charlotte. 1981. "Trends in Nonresponse Rates, 1952–1979." *Public Opinion Quarterly* 45:40–57.

U.S. Department of Health, Education, and Welfare. 1973. *Records, Computers, and the Rights of Citizens*. Washington, DC: Government Printing Office.

Westin, Alan F. 1967. *Privacy and Freedom*. New York: Atheneum.

———. 1981. *The Dimensions of Privacy*. New York: Garland.

———. 1990. *The Equifax Report on Consumers in the Information Age*. Atlanta: Equifax.

DIRECT AND INDIRECT ROUTES TO POLITICIZING PERSONAL EXPERIENCE

DOES KNOWLEDGE MAKE A DIFFERENCE?

DIANA C. MUTZ

Abstract This study examines the effects of issue-specific knowledge on the extent to which personal unemployment experiences influence presidential approval. The well informed are found to be more likely to connect personal experiences directly to political preferences, yet less likely to generalize from their own personal experiences in assessing the state of the nation as a whole. Since perceptions of the state of national conditions have a well-documented influence on presidential approval, simultaneously considering direct and indirect paths helps to resolve contradictory implications in past findings on the role of information in conditioning the politicization of personal economic experience.

The weak overall impact of self-interest on political preferences is widely known and well documented (for a review, see Sears and Funk 1990). Subsequent research in this vein has shifted toward locating conditions under which personal experiences are most likely to be politicized (see, e.g., Feldman 1985). The concepts used to locate groups most likely to rely on personal experiences have run the gamut from distinctions made on the basis of socioeconomic position (Weatherford 1983), habitual versus occasional voting (Cohen and Uhlaner 1991), levels of issue-specific knowledge (Conover, Feldman, and Knight 1986), and levels of news media use (Mutz 1992; Weatherford 1983).

Scholars generally concur about the distinction relevant to these comparisons: those with high levels of political knowledge, heavy use of mass media, high levels of political interest, and regular involvement in the political process are expected to differ in their decision-making

DIANA C. MUTZ is an assistant professor in the Department of Political Science and the School of Journalism and Mass Communication at the University of Wisconsin—Madison. The author would like to thank the Letters and Sciences Survey Center at the University of Wisconsin—Madison for making these data available and Dick Brody for his comments on an earlier draft of this study.

Public Opinion Quarterly Volume 57:483–502 © 1993 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/93/5704-0006$02.50

processes from those who are less knowledgeable, less educated, less interested, or less involved. Since these criteria are closely related, stable personal attributes, researchers are comparing roughly similar groups.

But the agreement ends here. Contradictory hypotheses have been proposed about which of these two groups utilizes personal experiences more and how their personal experience enters into their political judgments. One school of thought suggests that high levels of media use and political information should facilitate self-interested political attitudes (Sears et al. 1980, 1983). Another predicts that the effects of personal experience should be greater for the inattentive and poorly informed (Cohen and Uhlaner 1991; Conover, Feldman, and Knight 1986; Weatherford 1983). Neither side of this debate has produced compelling evidence. To complicate the research question further, in some cases the focus of inquiry is on the direct impact of personal experience on attitudes toward political actors; in others, it is on the indirect impact of personal experience by means of its effect on judgments about the state of the nation, which in turn influence attitudes toward political actors.

In this article I examine both the extent to which personal experience with unemployment directly influences presidential approval, and the extent to which unemployment experience indirectly affects presidential approval by means of its effect on assessment of national unemployment conditions. My purpose is to suggest a potential way of resolving a seemingly contradictory body of evidence.

Theoretical Bases for Contradictory Expectations

Two theoretical rationales have been offered to bolster the expectation that the well informed should be more likely to base their political preferences on their personal experiences. One theory suggests that well-informed people will be more sensitive to the personal costs and benefits associated with political outcomes. As Sigelman and Knight (1985) suggest, the more informed should be more likely to blame a president for problems over which he could fairly be expected to have some control. According to this logic, well-informed people should be more likely to recognize that the president is responsible for changes in unemployment and thus should be more likely to reward or punish the incumbent for their own personal employment problems.

A second theory suggests that the well informed should be more likely to politicize personal experiences because exposure to greater amounts of political information will make them more aware that their personal experiences are shared by many others. This, in turn, will make them more likely to reward or punish political actors. As Popkin explains it, "Information will affect my perception of whether a problem is mine alone or common to many" (1991, p. 11).

Some types of information might encourage the perception that unemployment is an individual rather than a collective problem, and thus discourage accountability. But unemployment news tends to be framed thematically, and thus draws attention to the collective nature of the problem (see Iyengar 1991).[1] The highly immediate, vivid, and accessible nature of personal unemployment experience, combined with an awareness of unemployment as a social problem, should produce greater political accountability. At a more general level, the greater degree of belief constraint among the well informed indirectly suggests that they should be more likely to make connections between personal and political judgments (e.g., Converse 1964; Sniderman, Brody, and Tetlock 1991).

Results from the few studies addressing the hypothesis that knowledge, political interest, or related concepts facilitate the effects of personal experience have been mixed. Sears and colleagues (1980) found that the strength of self-interested policy attitudes was not influenced by levels of political sophistication across four issue areas (see also Sears et al. 1983).[2] Abramowitz, Lanoue, and Ramesh, on the other hand, found that "voters who followed the presidential campaign closely were more likely to connect their personal financial situation with macroeconomic trends or government policies" (1988, p. 860). I have also found greater politicization of personal experiences among those who were heavy users of news media (Mutz, in press).

The contrary claim, that personal experiences will be more highly politicized by the poorly informed, stems from considering the "information costs" needed to acquire particular types of information. As Cohen and Uhlaner (1991) note, reliance on retrospective personal experience as a guide to voting was proposed largely as a low-cost alternative source of information for inattentive voters. According to this argument, more sophisticated voters will make use of more sophisticated types of information, while the less sophisticated who do not have relevant macroeconomic information will "default" to personal experiences as a source of information on which to base their evaluations of policies and politicians (Conover, Feldman, and Knight 1986, 1987; Weatherford 1983). As I suggest, in the absence of external information about broader social conditions, or the ability to comprehend or interpret it, "judgments will necessarily reflect more parochial concerns derived from personal experiences or interpersonal sources" (in press).

---

[1] Iyengar (1991) reports this finding for television news. Although this has not been a topic of empirical study, newspaper news seems even less likely to be episodic in the nature of its reporting since there is less emphasis on visuals and more space available for putting individual events in context.

[2] "Political sophistication" in this case was made up of education, political knowledge, and campaign interest, as well as four campaign media use items.

In studies of pocketbook voting and self-interested political preferences more generally, empirical evidence for the "default-source" theory is sparse and inconsistent. Conover, Feldman, and Knight (1986) found some indirect support for the hypothesis: the less informed relied on personal economic experiences in forming perceptions of inflation, while the more informed did not. A similar, though statistically insignificant, pattern was found for personal experience with unemployment. In both of the examples from their study, information weakened the extent to which personal experiences were used to make political evaluations, but only indirectly by weakening the impact of personal experience on perceptions of national conditions, which in turn influenced political attitudes. In a similar study, I (1992) found that those less knowledgeable about national unemployment conditions weighted personal experiences with unemployment more heavily in forming perceptions of national unemployment conditions.

Weatherford (1983) also found greater politicization of personal experience among the less informed. Moreover, he suggests that there is a trade-off between pocketbook and sociotropic voting—between reliance on personal experience and national referents—so that as reliance on one increases, the other decreases in strength. The well informed rely on national referents to form evaluations of political leaders, while those less informed rely on personal experiences. He found evidence suggesting that perceptions of prospective collective business conditions were weighted more heavily in evaluating economic performance among heavy print media users, while personal unemployment conditions were weighted more heavily by low print media users. This evidence is problematic, however, since comparing prospective perceptions of business conditions with retrospective personal unemployment experiences produces variations in both content and time frame, as well as personal versus collective emphasis. Nonetheless, some additional evidence partially corroborates his claim: Sears and colleagues (1980) found that the contribution of "symbolic" attitudes to policy opinions was greater among the politically sophisticated, but found no differences in the extent to which personal experiences were relied upon.[3]

---

[3] In this case, however, "symbolic" attitudes take on a quite different meaning. What Weatherford (1983) calls "symbolic" attitudes are judgments at the level of the collective, while Sears et al. (1980) use this same label to describe the impact of variables such as party identification, ideology, and racial tolerance.

MODELING DIRECT AND INDIRECT IMPACTS OF PERSONAL EXPERIENCE

If on first consideration this evidence seems inconsistent, upon closer examination it is less so. It is possible that high levels of information and political sophistication simultaneously increase the importance of personal experiences to political evaluations, yet decrease the extent to which personal experiences are used to form impressions of collective, national-level conditions. The Conover, Feldman, and Knight (1986) study assessed only the relationship between personal experience and judgments about collective conditions, that is, to what extent people generalized from personal experience to form national-level evaluations. The Weatherford (1983) study, on the other hand, assessed the importance of these contingencies on the extent to which personal experiences directly informed evaluations of political actors.[4]

In this study I hypothesize (1) that the well informed will be more likely to connect personal experiences directly to political preferences, yet (2) they will be less likely to generalize from their own personal experiences in assessing the state of the nation as a whole. Those well informed about macroeconomic conditions will see their personal experience as part of a larger phenomenon and thus link it to presidential support. But that same knowledge of national economic conditions will prevent them from relying on unrepresentative information based on personal experience in forming judgments about the nation as a whole. Since perceptions of the state of national conditions have a well-documented influence on presidential approval, the contradictory implications of direct and indirect effects of personal experience may account for the null and inconsistent findings in this research area.

In order to examine thoroughly the relationship between personal experience, perceptions of national conditions, and incumbent evaluations, I have also incorporated some possibilities not included in earlier studies. For example, the default source hypothesis maintains that personal experience serves as an easily accessible alternative information source in the absence of information on national conditions. But posing this question as a choice between only two alternatives is misleading. People might just as logically "default" to information about local economic conditions, their immediate social group, or the experiences of friends and acquaintances. Although it is impractical to incorporate all potential intermediate collectives, the extent to which people hear about unemployment through interpersonal interaction represents one level of collective judgment between personal experience and national experience, and thus takes into account a major source of information people may have about unemployment that is based on neither personal experience nor media-originated perceptions of national conditions.

---

[4] Still finer distinctions can be made between studies that have used general presidential job approval (e.g., Mutz 1992) and those using approval specific to the president's handling of a particular issue (Weatherford 1983).

In addition, both retrospective and prospective national economic judgments have been demonstrated to have important influences on attitudes toward political leaders (see, e.g., Fiorina 1981; Key 1966); however, it is strictly in the study of voting on the basis of retrospective perceptions that the issue of information sources has been emphasized. In retrospective voting models, the question of whether people reward or punish political leaders according to their own economic fortunes or according to their perceptions of the state of the nation calls for evaluating reliance on past personal experience as opposed to perceptions of others' collective experience based on information from external sources. When voting on the basis of prospective judgments, on the other hand, both personal and collective forecasts call for information outside the individual's life space. Researchers have found no differences between high- and low-knowledge groups in the weighting of personal experience in forming prospective assessments, nor did such a pattern emerge in this study; thus, the focus remains on retrospective perceptions. But personal experiences may still enter into evaluations of presidential performance by means of their influence on prospective national evaluations of unemployment. Thus any evaluation of the total impact of personal experiences must take into account effects that are mediated by prospective as well as retrospective national-level judgments.

Method

Data for my study come from a national telephone survey of approximately 1,200 respondents conducted by the University of Wisconsin Survey Center during the fall and winter of 1988–89.[5] Professionally trained, paid interviewers asked respondents a large set of parallel questions concerning personal and collective-level unemployment judgments. Telephone surveys are notorious for underrepresenting people at the low end of the socioeconomic spectrum, and this could be particularly problematic for a study involving unemployment since it is just such people who are especially susceptible to employment downturns (see Schlozman and Verba 1979). However, respondents in this survey turned out to be representative of the level of unemployment during that time period (see Appendix for survey details).

[5] Interviewing continued from September through January 1988, covering approximately a 5-month period.

The dependent variable used to examine the direct impact of personal experience on political preferences involves asking respondents to evaluate the president's overall performance on a 5-point scale.[6] For the examination of indirect effects, the dependent variable consists of retrospective assessment of national unemployment conditions.[7]

The independent variables in these analyses include personal- and national-level judgments about unemployment. Personal unemployment experience was assessed by asking whether the respondent or anyone in his or her immediate family had been unemployed or had had trouble finding a job in the past year.[8] Variables involving national-level judgments included retrospective and prospective judgments about unemployment in the United States since both represent indirect means by which personal experience may be politicized, and important direct influences on political judgments.[9]

---

6. "Overall do you approve or disapprove of the way Ronald Reagan is doing his job as President? Is that strongly approve/disapprove or somewhat approve/disapprove?" (Coded as 5-point scale; high = approve). This question was asked early in the half-hour survey so as to avoid artificially inflating the importance of personal or collective unemployment experience to incumbent evaluations (see Sears and Lau 1983). The same logic was used in ordering personal- and national-level judgments, with personal asked last so as not to inflate their importance to national evaluations. The order of other questions dealing with personal and national perceptions of unemployment might influence relationships among them, but it is unlikely to influence their interactions with knowledge since all were asked the questions in the same order.

7. "Would you say that over the past year people in the United States have had a harder time finding enough work, an easier time, or have things stayed about the same? Is that a little harder/easier or a lot harder/easier?" (Coded as 5-point scale; high = easier.)

8. "In the past year have you or anyone in your family been laid off or had trouble finding a job?" (Yes = 1, No = 2.) Although the unemployment rate was relatively low at this time (5.3 percent), including all those unemployed or having trouble finding work within the past year increased the number personally experiencing this problem to 8.2 percent. Extending this further to anyone in the respondent's immediate family increased the number of "Yes" answers to 20 percent, or 240 cases.

9. "How about people out of work during the coming 12 months—do you think that there will be more unemployment than now, less unemployment than now, or about the same?" (Coded as 3-point scale; high = less than now.) Because of their generally weaker relationship to evaluations of incumbents, prospective evaluations typically are placed to the left of retrospective evaluations in causal models and to the right of personal experiences (e.g., Conover, Feldman, and Knight 1986). Prospective evaluations at T_1 are expected to color retrospective evaluations at T_2; if expectations for improvement were high, people should be less likely to conclude retrospectively that things have improved. Nonetheless, in their panel study addressing the origins of retrospective and prospective judgments, Conover, Feldman, and Knight (1986, 1987) found that economic forecasts did not affect subsequent retrospective assessments; nor did retrospective assessments of economic conditions have much influence on economic forecasts. Whether considering simultaneous relationships or over-time panel correlations, retrospective and prospective national unemployment judgments have been found to be only modestly correlated ($r = .15$). Prospective judgments are especially likely to be sensitive to macroeconomic information from outside an individual's immediate life space (see Mutz 1992). Thus, despite potential simultaneity problems in estimating the relationship between prospective and retrospective judgments, prospective judgments are included as a control variable in these equations to prevent potential bias in estimat-

Since measures of collective unemployment conditions represent respondents' subjective assessments of national conditions, measures of researcher-defined personal experience are not entirely comparable. In this study, direct comparisons of personal and national judgments were facilitated by assessing subjective personal concern about unemployment with a question closely paralleling perceptions of national conditions. The measure was designed to tap those personally concerned about unemployment with or without actual personal experience.[10]

In addition, I included a variable indicating the amount of information regarding unemployment the respondent had received interpersonally. This measure is included to assess whether respondents are "defaulting" to information of this nature in the absence of abstract, national-level information instead of defaulting to personal experience-based judgments.[11]

With the exception of demographic and party identification variables, all indicators are coded so that the theoretically expected relationships are positive ones.[12]

DEFINING GROUPS FOR COMPARISON

Theories offered to account for differences in the extent to which people politicize personal experience vary not only in the anticipated direction of effect but also in specification of the characteristic that facilitates these differences. There is no single trait agreed upon as the central defining element distinguishing those more and less likely to politicize personal experiences. Weatherford (1983) chooses high and low levels of reliance on print news media for information about politics. Conover, Feldman, and Knight (1986, 1987) use measures of respondents' knowledge of the current inflation and unemployment

---

ing the relationship between retrospective judgments and the variables of theoretical interest. At worst, including prospective judgments overcontrols for the variance that otherwise might be accounted for by the effects of knowledge and its interactions with other variables.

[10] "How about you, or people in your own household . . . over the past year, have you been more worried about finding a job or keeping the one you have, less worried about these things, or have things stayed about the same? Is that a little more/less worried or a lot more/less worried?" (Coded as 5-point scale; high = less worried.)

[11] "How often do other people talk to you about their employment problems, that is, having trouble finding or keeping a job? Would you say they talk to you about job security or unemployment problems every day, 3 or 4 times a week, once or twice a week, or less often than that?" (Coded as 4-point scale; high = less often.)

[12] Demographic variables were coded as follows: party identification (Republican [2] or not [1]; Democrat [2] or not [1]); family income (6-point scale); age (coded in years); education (5-point scale); race (nonwhite [2] or white [1]); gender (recorded as female [2] or male [1]).

rates. Cohen and Uhlaner (1991) divide the electorate into habitual and irregular voters on the assumption that the preferences of habitual voters are based upon more information than those of irregular voters.

Although these indicators are no doubt correlated, respondents should be differentiated on the basis of the indicator that is closest to the hypothesized underlying cause. Defining groups by means of knowledge levels is often problematic in this regard, since such indicators are highly sensitive to the specific type of knowledge addressed. Widely criticized measures of "textbook knowledge" may tap something very different from questions addressing current issue-specific knowledge. In addition, to the extent that political preferences are formed through "on-line" evaluation, information that has influenced political preferences may no longer be evident at the time the preference is elicited (Lodge, McGraw, and Stroh 1989). However, since levels of political knowledge and interest tend to be stable personality characteristics that are highly correlated over time, this concern may not present serious problems in broad distinctions between those with high and low levels of knowledge.

For purposes of this study, issue-specific knowledge was chosen since most theories suggest that it is actual information about current unemployment conditions that should determine the processing strategy used. For example, the default source theory suggests that those without knowledge of national unemployment conditions should default to personal experiences in forming national-level judgments, precisely because they lack alternative information to inform that specific judgment.

Respondents were asked whether they knew the current unemployment rate in the United States; those within a full percentage point in either direction of the correct answer were deemed knowledgeable. The national unemployment rate had been at a flat 5.3 percent for over 6 months, perhaps making this question a bit easier than usual. Nonetheless, just under 30 percent of respondents fell into this range.[13]

Analysis

I will first assess the extent to which those with high and low levels of unemployment knowledge politicize their personal experiences by tying them directly to evaluations of presidential performance. These equations include perceptions of retrospective national unemployment to control for possible effects that flow through personal experience's effects on retrospective collective-level judgments. Next I examine the indirect effects of personal experiences on presidential evaluations by virtue of their influence on the formation of simple retrospective judgments about national unemployment. For each test, a dummy variable indicating knowledge level was coded so that the interaction represents the impact of the test variable in the indicated subgroup. The general models used to test hypotheses about direct and indirect effects are as follows:

---

[13] Only 6 percent of responses fell under 4.3 percent, while 64 percent of estimates were over 6.3 percent. As might be expected, Republicans were significantly more likely than Democrats or Independents to give low estimates of the unemployment rate ($r = -.09$, $p < .01$), while Democrats were more likely to give higher estimates ($r = .12$, $p < .01$), but these differences disappeared after controlling for demographic disparities between the two parties.

Direct effects:

Approval of president = a
$+ b_1$(personal experience)
$+ b_2$(personal concern)
$+ b_3$(retrospective perception of national unemployment)
$+ b_4$(prospective national unemployment concern)
$+ b_5$(interpersonal unemployment information)
$+ b_6$(unemployment knowledge)
$+ b_7$(unemployment knowledge × variable of interest)
$+ e$.

Indirect effects:

Retrospective perception of national unemployment = a
$+ b_1$(personal experience)
$+ b_2$(personal concern)
$+ b_3$(interpersonal unemployment information)
$+ b_4$(prospective national unemployment concern)
$+ b_5$(unemployment knowledge)
$+ b_6$(unemployment knowledge × variable of interest)
$+ e$.

Since these are cross-sectional data, it is difficult to disentangle short-term changes in the variables of theoretical interest from long-term factors influencing perceptions of economic conditions and support for incumbents. Here, as in most other studies examining this topic, control variables must substitute for a measured baseline by accounting for as much of the preexisting variation as possible. Demographic variables are likely to undercontrol for respondents' long-term political attitudes (Kiewiet 1983). Controlling for party identification overcontrols for long-term effects since these measures are subject to influence from short-term economic changes and political events (e.g., Page and Jones 1979; Rivers 1980; Shively 1980).

In the models described above I err on the conservative side by including dummy variables representing Republican and Democratic party identification as well as indicators of age, race, family income, gender, and education. Although this may attenuate the estimates somewhat, it greatly reduces the likelihood of reciprocal influences among variables and lends greater confidence to causal interpretations of the relationships that are sustained even with stringent controls.

The central focus in all of the models tested is the extent to which personal- or collective-level judgments are weighted differently in forming political preferences among the well informed as opposed to the poorly informed. These differences are tested in the equations described above by including an interaction between unemployment knowledge and the variable of interest. For example, to test the hypothesized interaction between knowledge and personal experience, a dummy variable was coded one for the more knowledgeable and zero for the less knowledgeable, so that the interaction between knowledge and personal experience represents the hypothesized additional impact of personal experiences among the well informed. Both personal- and collective-level variables were checked for potential interactions in order to assess the feasibility of Weatherford's (1983) idea that there is a trade-off between reliance on personal and national referents.

To insure that significant interactions did not result from multicollinearity among variables or from artificially constraining other coefficients in the pooled equations to be equal across the two subgroups, several precautions were taken. First, separate regressions were run within the two knowledge groups to identify overall patterns in the weighting of personal and collective judgments. If the confidence intervals surrounding the independent regression coefficients for the two groups did not overlap,[14] the key interactions were confirmed by including all respondents in a single equation incorporating a dummy variable designating high and low levels of knowledge and an interaction between knowledge and the variable of interest. Finally, I tested the increment to R^2 associated with the interaction terms to insure further that the interaction made a unique contribution to the model. The same procedure was repeated to identify significant interactions in the equation evaluating indirect effects.[15]

---

[14] See Greene (1990) for details on how this was determined.

[15] To facilitate comparisons with similar studies (e.g., Kiewiet 1983) results are presented using ordinary least squares regression. The major findings were the same when analyses were run using logit regression.

Results

DIRECT EFFECTS ON PRESIDENTIAL APPROVAL

Preliminary analyses revealed overall patterns similar to those found in other studies. The strongest single predictor of presidential performance aside from party identification was people's perception of national unemployment conditions. As in many other studies, personal unemployment experience was not related to support for the president. But personal unemployment concern and interpersonal information about unemployment were both significantly related to presidential approval. People who were more personally concerned than in the previous year and those who had heard through friends and acquaintances about employment problems were less likely to evaluate the president favorably. The main effect of knowledge of current unemployment conditions was to lower people's evaluations of Reagan's performance.

Table 1 illustrates the two significant interactions that emerged from examining the weighting of personal- and collective-level judgments in evaluating presidential performance. While personal experience with unemployment had no effect on evaluations of the president among the sample as a whole, the interaction between knowledge and personal experience indicates that personal experience did matter among the highly knowledgeable. For this group, a personal experience with unemployment lowered their already low evaluations of the president even more.

As indicated in table 1, retrospective perceptions of national conditions also had a greater influence on presidential approval among the more knowledgeable.[16] These same perceptions also had a significant main effect, but they had an additional impact among the knowledgeable. Although the size of the interaction coefficient is modest (b = .14), if one combines the two coefficients, the total impact of perceptions of national unemployment on presidential approval is .30 among the knowledgeable.

Thus far results conform to the expectations of the first hypothesis: personal experiences play a greater role in influencing political evaluations among the well informed. Although the size of the interaction coefficient is small, and the overall effects of personal experience are very modest, there does appear to be a difference in how these two

---

[16] As might be expected, party identification also mattered more in evaluating the incumbent among the knowledgeable group. Well-informed Republicans were particularly likely to evaluate Reagan favorably (interaction b = .42, $p < .05$), just as knowledgeable Democrats were more likely to evaluate him unfavorably (interaction b = $-.37$, $p < .05$).

Table 1. Effects of Personal- and Collective-Level Unemployment Perceptions on Approval of the President

Personal experience with unemployment	−.17 (.13)
Personal concern about unemployment	.12 (.04)**
Retrospective perceptions of national unemployment	.16 (.04)***
Prospective perceptions of national unemployment	.19 (.06)***
Interpersonal information about unemployment	.11 (.04)**
Unemployment knowledge	−1.51 (.47)**
Knowledge × personal experience	.26 (.11)*
Knowledge × retrospective national unemployment	.14 (.07)*
Party identification:	
Republican	.78 (.09)***
Democrat	−.61 (.09)***
Demographic controls:	
Age	−.00 (.00)
Gender	−.13 (.08)
Race	−.23 (.10)*
Income	.04 (.02)
Education	−.11 (.03)**
R^2	.33
N	1,062

NOTE.—The dependent variable in this equation is the 5-point presidential job approval scale (see Appendix). The knowledge dummy variable is coded 0 for low knowledge and 1 for higher knowledge so that the interactions represent the additional effects of the variable among the knowledgeable.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

groups make use of personal unemployment experience in evaluating the president. However, there is no evidence of a displacement or trade-off pattern; in fact, personal experience *and* perceptions of collective-level experience matter more among the well informed.

INDIRECT EFFECTS ON PRESIDENTIAL APPROVAL

To examine the indirect effects of personal experience, I regressed perceptions of national unemployment conditions on the remaining independent variables, including the indicator of unemployment knowledge and an interaction between knowledge and personal experience. In order to test the hypothesis that personal-level variables will matter more for the less informed, the coding of the dummy variable in the interaction was reversed so that the interaction represented the hypothesized additional impact of personal experiences among the less knowledgeable.

The results in table 2 confirm the second hypothesis: personal experiences influenced national unemployment perceptions more among those with low knowledge. Personal experience did not have a significant effect in the sample as a whole, but specifically among the less knowledgeable, having a personal experience with unemployment lowered evaluations of national unemployment conditions by almost half of a point on the 5-point scale. If one subtracts the negative main effect ($b = -.16$), this still leaves a net impact of .29.

The main effect of knowing the unemployment rate was to make people more pessimistic in their assessments of national unemployment conditions, just as the knowledgeable were more negative in their evaluations of the president. Those with low unemployment knowledge were more favorable in their assessments of unemployment regardless of personal unemployment experience. But those who had had

Table 2. Effects of Personal- and Collective-Level Unemployment Perceptions on Assessments of National Unemployment Conditions

Personal experience with unemployment	−.16 (.17)
Personal concern about unemployment	.19 (.03)***
Prospective perceptions of national unemployment	.18 (.05)***
Interpersonal information about unemployment	.14 (.03)***
Unemployment knowledge	−1.01 (.36)**
Knowledge × personal experience[a]	.45 (.19)*
Party identification:	
Republican	.08 (.08)
Democrat	−.24 (.08)**
Demographic controls:	
Age	−.00 (.00)
Gender	−.21 (.07)**
Race	−.14 (.09)
Income	.08 (.02)***
Education	−.00 (.03)
R^2	.20
N	1,062

Note.—The dependent variable in this equation is the 5-point scale representing retrospective perceptions of national unemployment (see Appendix).

[a] The knowledge dummy variable in the interaction has been recoded 1 for low knowledge and 0 for high knowledge so that the interaction represents the additional effects of personal experience among the less knowledgeable.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

no personal experience with unemployment were even more positive in their assessments. As predicted, people with high levels of knowledge did not use their personal experiences in forming evaluations of national conditions. But those without knowledge of the current employment situation relied instead on their own personal unemployment experience.

Both groups, however, relied on their personal concerns about unemployment and their interpersonally obtained information about others' employment problems: the less knowledgeable were no more likely to "default" to these types of measures than were the more knowledgeable. Prospective assessments also made a significant contribution to explaining perceptions of national conditions; those perceiving a positive overall trend were also more generous in their assessment of recent changes. And finally, partisan rationalization also played a role in that Democrats were systematically more pessimistic about national unemployment conditions.

LIMITATIONS OF THE STUDY

Although the pattern of results observed in this study makes a great deal of sense in terms of the hypothesized relationships, some caution should be exercised in comparing the contributions of personal experience and those of variables such as personal concern, interpersonal unemployment discussion, and prospective national unemployment. The potential for reciprocal influence is much less with measures of objective personal experience than for the variables that incorporate greater personal judgment.

Although possible reciprocal influences are always present to some degree, two factors suggest that the risk is minimal in this study. First, other studies explicitly addressing these problems using two-stage least squares analyses or panel data have found no support for the reciprocal influences of concern here (see, e.g., Conover, Feldman, and Knight 1986, 1987; Mutz 1992). In addition, including a stringent battery of control variables greatly reduces the potential for simultaneity (see Kiewiet 1983).

As with most survey-based findings, caution should also be exercised in generalizing results from a single point in time, particularly one in which unemployment levels were relatively low. Finally, it is also possible that nonresponse damaged the representativeness of the survey sample in ways that have influenced the reported findings. This seems unlikely, however, since the analyses do not attempt to describe a population, but rather seek to compare the strength of relationships among variables between subgroups.[17]

[17] It is possible that underrepresentation of blacks and Hispanics contributes to under-

Discussion

Does knowledge increase or decrease the translation of personal concerns into political preferences? Based on examinations of the impact of personal unemployment experience, the answer to this question is that knowledge may do both. It may simultaneously increase the extent to which people who are aware of unemployment conditions hold the president responsible for their personal unemployment experience and decrease the strength of this relationship by virtue of the extent to which national conditions are mere projections of personal experience.

What is perhaps most striking about these results is the extent to which personal-level variables enter into evaluations of the president—even when collapsing across knowledge groups, and despite the fact that unemployment was relatively low. Previous research suggests that among economic problems, unemployment is particularly difficult for people to link to political actors (Kinder and Mebane 1982, p. 150). The low level of unemployment and low likelihood of politicizing personal unemployment concerns make the significance of the personal concern measures all the more noteworthy. These findings suggest that people's failure to connect the personal and the political is to some extent a methodological artifact of differences in how personal- and national-level judgments are typically measured.

Personal experiences and perceptions of national-level conditions are only two points along a continuum running from an individual's immediate life space to his or her perceptions of successively larger collectives. By pitting referents at opposite ends of this continuum against one another as predictors of political attitudes, some studies of pocketbook and sociotropic voting have implicitly, if unintentionally, furthered the idea that relying heavily on one referent means relying less on another. Instead, my findings jibe well with the long-standing notion of more and closer relationships between constructs among the well informed. In this study I found no support for the notion of a trade-off between reliance on personal- and collective-level perceptions. Those with knowledge of current unemployment were more likely to use both personal experiences and collective-level judgments of unemployment in forming their evaluations of the president.

These results also shed light on the extent to which personal experiences truly serve as "low-cost" or "default" sources of political information. As social scientists increasingly focus on how people with

---

estimating the extent to which personal experiences are politicized since group identities sometimes aid people in connecting personal and political interests. On the other hand, underrepresenting those with low levels of education may inhibit the politicization of personal experience, so the consequences of underrepresentation are unclear.

limited information form political opinions (see, e.g., Ferejohn and Kuklinski 1990; Popkin 1991; Sniderman, Brody, and Tetlock 1991), the issue of information costs has drawn additional attention. The lack of effort required to obtain information from personal experience generally relegates it to the low cost category. People are assumed to rely on highly accessible personal life experience as a substitute for more abstract, high-cost political information. But connecting one's personal experiences directly to political preferences is notoriously difficult. Moreover, the costs of obtaining information may be overshadowed by the costs of interpreting it.

The broader implications of this study are somewhat discouraging. In one sense the results may seem reassuring since there is a means by which both the knowledgeable and the less knowledgeable may politicize their personal unemployment experience and hold political leaders accountable. To the extent that less knowledgeable people generalize from their more immediate environments in forming perceptions of national conditions, sociotropic politics becomes simply an extension of pocketbook politics. But overall, personal experiences matter very little, particularly among those with little knowledge of macroeconomic conditions. Even if one combines the direct and indirect impacts of personal experience on presidential approval, the total amount of influence among those with high knowledge is still four times its impact among the less knowledgeable. In addition, personal concern has almost twice as much total influence among those with high unemployment knowledge.[18]

The findings of this study confirm the long-standing notion that well-informed citizens will be better equipped to hold politicians accountable for the effects of their economic policies. From a normative perspective, the problem is that the well informed are not those most likely to be personally experiencing the effects of unemployment or, for that matter, many other social problems. For example, in this study women and minorities were only half as likely to be knowledgeable about unemployment as were men and nonminorities,[19] but these same groups are more likely to be experiencing the consequences of unemployment personally.[20] Likewise, those with incomes under $40,000 were only half as likely to have unemployment knowledge as those

---

[18] These figures were based on combining the direct impact of personal experience (or concern) with its indirect effects that flow through retrospective or prospective national perceptions.

[19] Chi-square values comparing the percentage knowledgeable by gender and race were 42.55 ($p < .001$) and 15.22 ($p < .001$), respectively.

[20] The economic problems of women and minorities have been well documented elsewhere, and this same pattern holds with respect to the incidence of unemployment among women and minorities in this sample (χ^2 values were 10.53 [$p < .01$] and 16.88 [$p < .001$], respectively).

with incomes over \$40,000.[21] In short, a lack of macroeconomic information insulates politicians from the very segments of the population most likely to be adversely affected by employment downturns. Inequalities in the distribution of information, coupled with depressed voting turnout among those with low socioeconomic status, combine to severely limit presidential accountability for personal unemployment experience.

Appendix

Survey Sampling Procedures

The Letters and Science Survey Center (LSSC) at the University of Wisconsin—Madison runs a continuous half-hour national telephone survey. The center's sampling procedure begins with a sample of telephone numbers representative of currently working residential telephone numbers in the continental United States purchased from Nielsen Media Research.

One person is selected at random from among the adult (age 18 or older) members of the sample household. Each sample number is called up to 10 times, using a "day of the week" calling strategy. Each day's interviews constitute a random sample of the population on that day. This requirement means that it is necessary to deal with the problem of nonresponse in a special way, using a procedure first suggested by Kish and Hess (1959) and elaborated upon by Madow, Hyman, and Jessen (1961). The procedure depends on replacing current not-at-homes with not-at-homes saved from previous sample draws. Because the kinds of people not home on one day of the week may be different from those not at home on another, the replacement scheme is day-of-the-week specific.

The LSSC's response rate is approximately 52 percent if one includes both refusals and unresolved numbers in the denominator. This is considerably lower than the Center for Political Studies and the National Survey of Families and Households; however, distributions of respondents are quite similar with respect to major demographic variables including household size, region, age, sex, household income, marital status, religious preference, unemployment, and church attendance. Differences result from the fact that LSSC is a telephone survey that consequently underrepresents minorities and the less educated.

References

Abramowitz, Alan I., David J. Lanoue, and Subha Ramesh. 1988. "Economic Conditions, Causal Attributions, and Political Evaluations in the 1984 Presidential Election." *Journal of Politics* 50:848–63.

Cohen, Linda R., and Carole J. Uhlaner. 1991. "Participation, Closeness and

---

[21] Chi-square = 40.62 ($p < .001$).

Economic Voting in the 1960 Presidential Election." Paper presented at the annual meeting of the American Political Science Association, Washington, DC.

Conover, Pamela J., Stanley Feldman, and Kathleen Knight. 1986. "Judging Inflation and Unemployment: The Origins of Retrospective Evaluations." *Journal of Politics* 48:565–88.

———. 1987. "The Personal and Political Underpinnings of Economic Forecasts." *American Journal of Political Science* 31:559–83.

Converse, Phillip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. D. Apter, pp. 206–56. New York: Free Press.

Feldman, Stanley. 1985. "Economic Self-interest and Political Behavior." *American Journal of Political Science* 26:446–66.

Ferejohn, John A., and James H. Kuklinski. 1990. *Information and Democratic Processes*. Urbana: University of Illinois Press.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Greene, William H. 1990. *Econometric Analysis*. New York: Macmillan.

Iyengar, Shanto. 1991. *Is Anyone Responsible? How Television Frames Political Issues*. Chicago: University of Chicago Press.

Key, Valdimer O. 1966. *The Responsible Electorate*. Cambridge, MA: Harvard University Press.

Kiewiet, D. Roderick. 1983. *Macroeconomic and Micropolitics: The Electoral Effects of Economic Issues*. Chicago: University of Chicago Press.

Kinder, Donald R., and Walter R. Mebane, Jr. 1982. "Politics and Economics in Everyday Life." In *The Political Process and Economic Change*, ed. K. R. Monroe, pp. 141–80. New York: Agathon.

Kish, Leslie, and Irene Hess. 1959. "A Replacement Procedure for Reducing the Bias of Nonresponse." *American Statistician* 13(4):17–19.

Lodge, Milton, Kathleen M. McGraw, and Patrick Stroh. 1989. "An Impression-driven Model of Candidate Evaluation." *American Political Science Review* 83:399–419.

Madow, William G., Herbert H. Hyman, and Raymond J. Jessen. 1961. "Evaluation of Statistical Methods Used in Obtaining Broadcast Ratings." *House Report No. 193*, 87th Congress, 1st Session. Washington, DC: Government Printing Office.

Mutz, Diana C. In press. "Contextualizing Personal Experience: The Role of Mass Media." *Journal of Politics*.

———. 1992. "Mass Media and the Depoliticization of Personal Experience." *American Journal of Political Science* 36:486–507.

Page, Benjamin I., and Calvin C. Jones. 1979. "Reciprocal Effects of Policy Preferences, Party Loyalties, and the Vote." *American Political Science Review* 73:1071–89.

Popkin, Samuel L. 1991. *The Reasoning Voter*. Chicago: University of Chicago Press.

Rivers, Douglas. 1980. "The Dynamics of Party Support in the American Electorate, 1952–1976." Paper presented at the annual meeting of the American Political Science Association, Chicago.

Schlozman, Kay, and Sidney Verba. 1979. *Injury to Insult: Unemployment, Class and Political Response*. Cambridge, MA: Harvard University Press.

Sears, David O., and Carolyn L. Funk. 1990. "Self-interest in Americans' Political Opinions." In *Beyond Self-interest*, ed. J. J. Mansbridge, pp. 147–70. Chicago: University of Chicago Press.

Sears, David O., and Richard R. Lau. 1983. "Inducing Apparently Self-interested Political Preferences." *American Journal of Political Science* 27:222–52.

Sears, David O., Richard R. Laa, Tom R. Tyler, and Harris M. Allen, Jr. 1980. "Self-interest vs. Symbolic Politics in Policy Attitudes and Presidential Voting." *American Political Science Review* 74:670–84.

Sears, David O., L. Steck, Richard R. Lau, and M. T. Gahart. 1983. "Attitudes of the Post-Vietnam Generation toward the Draft and American Military Policy." Paper presented at the annual meeting of the International Society of Political Psychology, Oxford.

Shively, W. Phillips. 1980. "The Nature of Party Identification: A Review of Recent Developments." In *The Electorate Reconsidered*, ed. J. C. Pierce and J. L. Sullivan, pp. 219–36. Beverly Hills, CA: Sage.

Sigelman, Lee, and Kathleen Knight. 1985. "Public Opinion and Presidential Responsibility for the Economy: Understanding Personalization." *Political Behavior* 7:167–91.

Sniderman, Paul M., Richard A. Brody, and Philip E. Tetlock. 1991. *Reasoning and Choice: Explorations in Political Psychology*. New York: Cambridge University Press.

Weatherford, M. Steven. 1983. "Economic Voting and the 'Symbolic Politics' Argument: A Reinterpretation and Synthesis." *American Political Science Review* 77:158–74.

MULTIPLE-MEASURE ASSESSMENT OF PARTY IDENTIFICATION

DONALD PHILIP GREEN
ERIC SCHICKLER

Abstract Recent research has focused attention on the importance of accounting for measurement error in party identification when modeling the stability of partisanship and the determinants of the vote. Measurement error estimates have in the past been based on a single measure of partisanship observed at multiple points in time, a test-retest methodology that requires fairly strong assumptions about the character of change over time. This article assesses the reliability of the Michigan party identification scale using multiple measures of partisanship at a single point in time. Our data not only corroborate previous test-retest results but also suggest that the accuracy with which partisanship predicts candidate preferences can be enhanced using multiple measures. One measure in particular, a labeled 7-point self-placement continuum, is found to hold significant potential to supplement and illuminate the Michigan scale.

For decades the conceptualization and measurement of party identification have received considerable attention (cf. Budge, Crewe, and Farlie 1976; Campbell et al. 1960; Converse and Pierce 1985; Dennis 1988; Green and Palmquist 1990; Miller 1991). At the same time, however, the way in which surveys have assessed party identification has remained more or less uniform. The vast majority of national surveys conducted over the past 30 years have employed either the Michigan or the Gallup party identification measures (see Appendix for question wording), with private polling firms often using the latter and academic survey organizations leaning heavily toward the former. While uniformity of method is of considerable convenience to those who wish to

DONALD PHILIP GREEN is associate professor of political science at Yale University. ERIC SCHICKLER is a graduate student in political science at Yale University. This research was supported by grants from the Institution for Social and Policy Studies at Yale, 111 Prospect St., New Haven, CT 06520. The authors are grateful to the Inter-University Consortium for Political and Social Research and the Times-Mirror Organization for supplying the data used in this article. Thanks to Paul Abramson, Joel Fetzer, and Brad Palmquist for their comments on an earlier draft.

Public Opinion Quarterly Volume 57:503–535 © 1993 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/93/5704-0005$02.50

track responses to the same question over a long span of time, it imposes certain limitations on other research agendas. In particular, it hampers investigation of the statistical reliability of the party identification measure.

Reliability assessment in this case is more than an arcane technical endeavor; it lies at the heart of the debate about the malleability of party identification. The imputed reliability of the party identification scale directly affects one's impression of the dynamics of party affiliation. Analyses that assume party identification to be measured perfectly have found it to be unstable over time and responsive to short-term political forces such as candidate evaluations (Page and Jones 1979), retrospective performance evaluations (Brody and Rothenberg 1988; Fiorina 1981), ideological proximity to the parties (Franklin and Jackson 1983; Jackson 1975), and past voting behavior (Markus and Converse 1979). In contrast, analyses that have taken into account the finding that 10–15 percent of the variance of the 7-point Michigan party identification scale is random noise have found partisanship to be highly stable over time (Achen 1975; Asher 1974; Green and Palmquist 1990, and in press; Krosnick 1991; Palmquist and Green 1992) and unresponsive to the aforementioned short-term forces (Green 1990; Green and Palmquist 1990; Schickler and Green 1993).

Although reliability assessments by different authors have tended to produce similar results across a range of data sets, the way in which reliability has been assessed is not above criticism. In every case of which we are aware, reliability estimation has relied on techniques designed for panel data in which a single indicator of party identification is repeated at different points in time (Heise 1969; Wiley and Wiley 1970, 1974). As noted by Palmquist and Green (1992), Sullivan and Feldman (1979), and others, this test-retest method rests on strong assumptions about the character of the causal process, measurement error, and structural change. In particular, it is assumed that the errors associated with a measure at any given wave are independent of errors that occur at any other point in time. It is also assumed that the factors that produce change in partisanship are independent over time. Finally, one must assume that measurement errors are independent of the factors that induce true change in partisanship (for an extensive review of these assumptions, see Palmquist and Green [1992]).

These are not necessarily invalid assumptions, but they are sufficiently strong to warrant an investigation of reliability that does not rely on test-retest techniques.[1] The purpose of this article is to assess the reliability of the Michigan party identification measure using an alternative approach, one that involves multiple measures of partisanship at a single point in time. The advantage of this approach is that it does not require us to assume anything about the character of change in partisanship over time.[2] Using a variety of different measures of party identification appearing in two surveys, the December 1973 National Opinion Research Center (NORC) Amalgam Survey and the Times-Mirror Spring 1990 Political Update, we obtain results that are broadly consistent with earlier analyses that utilize test-retest methods.

---

[1] As Palmquist and Green (1992) point out, test-retest methods using single indicators offer certain advantages. The practical demands of survey design and space limitations are reduced, inasmuch as the researcher need only come up with one valid indicator of

In addition to facilitating structural equation modeling with latent variables, the inclusion of multiple indicators of partisanship has potential implications for prediction and survey design. How can survey instruments be augmented so that the vote is better predicted? To address this question, we assess the reliability of the supplementary measures of party identification. Of the four auxiliary measures of long-term party identification we consider, one is sufficiently reliable to enhance the predictive power of a regression model that includes only the standard Michigan party identification scale. Since this measure, a labeled 7-point self-placement scale, is relatively easy to administer in personal interviews, we recommend its adoption in face-to-face voter surveys.

Conceptualization and Measurement of Party Identification

Party identification is the central theoretical component of the more general construct of partisanship. While partisanship connotes such dimensions as group membership, behavioral expressions of commitment, psychological closeness, policy agreement, and electoral support, the concept of party identification focuses on the individual's self-definition (Green and Palmquist 1990; Campbell et al. 1960, 1986). As such, a valid measure of party identification requires that the respondent identify himself or herself in relation to the political parties. Measures of affect toward the parties are not valid measures of party identification; a valid measure of an individual's attachment to a party

---

the latent construct. In addition, the administration of just one measure at a given time point avoids the sort of priming and recency effects that might cause nonrandom errors across multiple measures, a concern expressed by Bishop, Tuchfarber, and Smith (1992).

[2] The disadvantage is that measurement errors associated with one question may be correlated with errors associated with other measures employed in the same survey. This problem can be minimized by placing different measures in different parts of the survey and changing the response formats offered. To their credit, the designers of the surveys used here made an effort to take both precautionary steps.

must distinguish his or her momentary likes and dislikes from "attributes felt to be part of his or her persona, or definition of the social self" (Converse and Pierce 1985, p. 144). Phrases such as "think of yourself," "call yourself," "place yourself," or "regard yourself" tap this self-concept, while feeling thermometers or questions that elicit prospective performance evaluations do not.

A related point is that identification is the facet of partisanship that is most likely to represent an enduring, underlying trait. Noting that the stability of partisanship over time is what lends the construct its explanatory power, Converse and Pierce (1985) observe that the standard Michigan partisanship question is particularly suited to this task because it taps self-identity with an extended time horizon. For unlike the wording used by the Gallup organization, which asks respondents about how they think of themselves "in politics as of today," the Michigan measure is prefaced with the phrase "generally speaking" or "in general." A good deal of recent scholarship has focused on the extent to which this difference in wording actually affects the over-time stability of partisanship, with conflicting results (Abramson and Ostrom 1991, 1992, 1993; Bishop, Tuchfarber, and Smith 1992; MacKuen, Erikson, and Stimson 1989, 1992). Common to both sides of this debate, however, is a recognition that identification is the key construct of interest.

Many surveys have included measures of different aspects of partisanship, ranging from identification to electoral support, but only a small handful have made serious efforts to measure party identification redundantly (see Abramson and Ostrom 1993). Here, we focus our attention on two such surveys, the December 1973 NORC Amalgam Survey and the Times-Mirror Spring 1990 Political Update.[3] The NORC survey used a national multistage area probability sample. At the block level, quota sampling was used. The 1,489 respondents were interviewed in person. The survey was designed in large part to test the effects of question wording. As a result, after the first several questions, respondents were randomly assigned to one of three groups.

---

[3] We elected not to analyze the Political Action Panel Study (Barnes et al. 1988), which contains two measures of partisanship, the Michigan scale and a self-placement scale on which the respondent states to which party he or she feels "closer." Although the "closeness" question is similar to the second branch of the Michigan scale (directed toward self-identified independents), it is not clear whether it validly measures party identification when not prefaced with a trunk question that specifically calls attention to sense of identity. Also unsuitable were the items available in the 1980 National Election Study (NES), which are either invalid measures of party identification (e.g., party support) or illuminate only specific ranges of the partisan spectrum, such as independence (Dennis 1988). Finally, we chose not to analyze the 1987 NES Pilot Study (Niemi, Reed, and Weisberg 1991) because the items it contains are less diverse than the NORC study we examine here. It should be noted that where the 1987 NES and 1973 NORC studies overlap, the results turn out to be exremely similar.

The three groups were asked identical first and third questions about their party identification, but each group was asked a different second question. The virtue of this design is that any one set of respondents was confronted with only three party identification questions dispersed throughout a lengthy survey, yet data from five measures of partisanship were obtained in all. In the Times-Mirror survey, 3,004 respondents were interviewed in person. The sample used for the survey was a multistage probability sample, consisting of 100 primary sampling units and 1,600 secondary sampling units. Interviewing was conducted in 193 sampling locations. The Times-Mirror study included numerous innovative items designed to gauge various dimensions of partisanship, although only two of the items meet the criteria for the valid measurement of party identification discussed above.

The NORC survey includes five measures of party identification. The first is the standard 7-point Michigan scale. The second item asks the respondent to place him- or herself on a fully labeled 7-point scale, with strong Democrats and strong Republicans on either end and independents in the middle (cf. Krosnick and Berent 1993). We refer to this item as the *7-point self-placement scale*. The third item, referred to as the *forced-pair comparison item*, is identical to the first part of the branched Michigan question but requires the respondent to volunteer independent partisanship. The fourth item contains an even more explicit long-term prompt than the Michigan item: respondents are asked how they "basically" think of themselves, regardless of recent voting behavior or future vote intentions. We refer to this as the *basic self-regard item*. The final item uses the standard Gallup "politics-as-of-today" question wording but does not probe the strength of partisanship or attempt to identify leaning partisans. We refer to this item as the *3-point Gallup question*. Unlike the other items in the NORC survey, the prompt in this item is explicitly short-term.

The Times-Mirror survey includes two measures of party identification. The first is a 7-point scale using the Gallup politics-as-of-today prompt. The second is the difference between the respondent's view of the suitability for him or her of the descriptions "a Republican" and "a Democrat." This party difference score, which arguably taps a longer-term self-conception, is a 19-point scale that ranges from −9 to +9.

Comparing Alternative Measures of Party Identification

UNIVARIATE STATISTICS

The marginal distributions of the seven measures of partisanship are presented in tables 1–4. Looking first at the five NORC questions, we

Table 1. Response Distribution of Michigan 7-Point Party Identification and 7-Point Self-Placement Scale: NORC

Identification	Michigan (%)	Self-Placement (%)
Strong Democrat	18.4	13.8
Weak Democrat	19.9	. . .
Prefer Democrats	. . .	14.8
Democrat, strength unclear	1.7	. . .
Leaning Democrat	13.5	18.9
Independent	9.4	22.7
Leaning Republican	7.5	13.0
Republican, strength unclear	1.1	. . .
Prefer Republicans	. . .	8.7
Weak Republican	12.4	. . .
Strong Republican	8.1	4.1
Other party	3.5	. . .
Don't know/refused	4.5	4.0
Total	100.0	100.0
N	1,489	507

Table 2. Response Distribution of Forced-Pair Comparison, Basic Self-Regard, and Gallup 3-Point Scale: NORC

Identification	Forced Pair (%)	Basic Self-Regard (%)	Gallup (%)
Democrat	47.9	43.8	40.5
Republican	23.3	19.3	21.0
Independent		27.8	33.4
Independent (volunteered)	18.7		
Other (volunteered)	3.2	1.8	2.1
Don't know/refused	6.8	7.3	3.0
Total	99.9	100.0	100.0
N	497	493	1,489

Table 3. Response Distribution of Gallup 7-Point Scale: Times-Mirror 1990 Study

Identification	Percentage
Strong Democrat	17.3
Weak Democrat	18.5
Lean Democrat	11.9
Independent	11.3
Lean Republican	12.0
Weak Republican	15.2
Strong Republican	12.3
Refused or not applicable	1.6
Total	100.1
N	3,004

find that the responses are fairly consistent across the Michigan, Gallup, and basic self-regard measures. The contrast in wording between Michigan and Gallup ("generally speaking" as opposed to "in politics as of today") seems to have had little effect on the marginal distribution of responses (see Bishop, Tuchfarber, and Smith [1992] for a similar result; and Abramson and Ostrom [1993] for a somewhat contrasting finding). When asked to ignore their past voting tendencies, respondents in 1973 tended to report greater Democratic affiliation. Even so, erasing the partisan coloration of two recent Republican presidential victories had only modest effects on respondents' reported party affiliation. Pooling the three items together—the 3-point Michigan classification, the Gallup measure, and the basic self-regard item—we find no support for the hypothesis that differences in wording affected the frequency of responses. A χ^2 (4) test of independence yields a statistic of 5.62, well below the critical value needed to reject the null hypothesis of no differences due to question wording ($\alpha = .05$).

Not surprisingly, the proportion of the sample identifying with the two major parties goes up significantly when the trunk question of the Michigan measure omits reference to independents. If anything, the surprise is that even when independent identification is a response option that must be volunteered, it still accounts for 19 percent of the sample. In contrast, when respondents are presented with a visual 7-point continuum on which they must place themselves, the availability of a middle option attracts 23 percent of the sample, more than twice the proportion identifying as pure independents on the Michigan scale. Evidently, as others (Barnes et al. 1988; Keith et al. 1992; Ken-

Table 4. Response Distribution of Party Difference Score: Times-Mirror 1990 Study

Party Difference Score	Percentage
−9	18.5
−8	2.8
−7	3.4
−6	3.4
−5	3.2
−4	4.2
−3	2.0
−2	3.4
−1	2.5
0	17.4
1	2.9
2	3.3
3	2.9
4	3.7
5	3.0
6	2.8
7	3.2
8	2.0
9	9.3
Don't know (for Republican description)	.5
Don't know (for Democratic description)	2.9
Not applicable	2.8
Total	100.1
N	3,004

Note.—See Appendix for scoring.

ney and Rice 1988) have suspected, the ratio of partisans to independents is a malleable quantity, sensitive to the format of survey items.

One further point concerning the NORC data deserves mention. Each respondent addressed three party identification questions, beginning with the branched Michigan scale and ending with the trunk of the Gallup scale. In between, respondents were assigned to one of three experimental questions. Thus, the Gallup question may be regarded as a posttest in a priming experiment, where the self-placement scale, the forced-pair question, and the basic self-regard questions serve as the primes. Interestingly, there is no evidence of any priming effect on either the mean (χ^2 [4] = 8.8, $p > .05$) or variance (F[2, 1,272] = 1.48, $p > .05$) of responses to the Gallup item. While this finding, by itself, does not show that priming effects are of no concern in multiple-measure assessments of party identification, it does at least indicate that when spaced a few minutes apart, consecutive measures do not bleed into one another.

Turning to the Times-Mirror data, we consider the 19-point scale constructed by taking the difference between "self-regard" ratings of the two parties. Like the Michigan or Gallup 7-point scales, the party difference score is nonnormal in appearance, with clusters at both endpoints. Despite the length of the scale, the endpoints account for 28 percent of the valid responses. In addition, like the self-placement scale examined above, approximately one-fifth of the sample ends up at the midpoint. Thus, while the many gradations of partisanship offered by the scale make it a potentially attractive item, the pattern of responses suggest that it fails to overcome some of the limitations of other measures.

BIVARIATE RELATIONSHIPS

When one cross-tabulates responses to the Michigan scale with answers to the alternative measures used in the NORC survey, the results indicate quite clearly that the standard party identification measure is not an errorless indicator of partisanship (see tables 5–8). Response variation is particularly evident in the more articulated scales, such as self-placement, where an occasional respondent identified as a strong partisan by the Michigan measure surfaces in the middle of the partisan continuum. But even when the 7-point scale is arrayed against a crude 3-point classification, "party switching" occurs. For example, when offered a forced-pair choice between the Democratic and Republican parties, three of the 198 respondents dubbed Democrats by the Michigan measure called themselves Republicans. Conversely, four of 94 ostensible Republicans called themselves Democrats. Thus, the total rate of interparty conversion is 7/292 (or 2.4 percent), a rate that exceeds what one typically observes in panel studies where respondents are interviewed several months apart (Green and Palmquist, in press). Somewhat less interparty conversion is evident when the comparison item is the 3-point Gallup measure, which furnishes a middle response option. Nevertheless, as Nie, Verba, and Petrocik (1979) note, 2.0 percent of the 917 Michigan partisans underwent "conversions."[4]

---

[4] The corresponding rates for the basic self-regard item and the self-placement scale are 1.9 percent (6/322) and 1.0 percent (3/309). The rate for the latter is depressed due to the relative paucity of respondents choosing to call themselves strong or weak partisans when presented with a 7-point continuum.

Table 5. Cross-Tabulation of Michigan 7-Point Party Identification and 7-Point Self-Placement

Self-Placement 7-Point Scale	Michigan Identification											
	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	Total
Strong Democrat	58	6	2	0	0	0	0	0	0	2	2	70
Weak Democrat	22	36	6	4	0	0	0	2	0	2	3	75
Lean Democrat	8	40	1	37	2	2	0	3	0	2	1	96
Independent	4	2	1	29	36	17	2	9	1	8	6	115
Lean Republican	0	0	0	0	3	23	4	27	7	2	0	66
Weak Republican	0	1	0	0	0	1	1	21	20	0	0	44
Strong Republican	0	0	0	0	0	0	2	2	17	0	0	21
Other/don't know	0	1	0	0	3	1	0	2	1	7	5	20
Total	92	86	10	70	44	44	9	66	46	23	17	507

Table 6. Cross-Tabulation of Michigan 7-Point Party Identification and Forced-Pair Comparison

Forced-Pair Comparison (Volunteer "Independent")	Michigan Identification										
	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	Total
Democrat	75	102	7	34	5	1	0	4	0	4	6	238
Independent (volunteered)	0	4	0	28	39	15	0	0	0	1	6	93
Republican	0	3	0	1	2	18	4	45	41	1	1	116
Other (volunteered)	0	0	0	0	5	0	0	0	0	10	1	16
Don't know/refused	2	5	0	2	6	1	0	0	0	1	17	34
Total	77	114	7	65	57	35	4	49	41	17	31	497

Table 7. Cross-Tabulation of Michigan 7-Point Party Identification and Basic Self-Regard

	Michigan Identification											
Basic Self-Regard	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	Total
Democrat	98	82	7	17	0	0	0	3	1	2	6	216
Independent	5	11	0	49	33	23	0	9	1	2	4	137
Republican	0	2	0	0	0	7	3	53	29	1	0	95
Other (volunteered)	0	0	0	0	1	0	0	2	0	6	0	9
Don't know/refused	2	7	1	0	6	4	0	3	3	1	9	36
Total	105	102	8	66	40	34	3	70	34	12	19	493

Table 8. Cross-Tabulation of Michigan 7-Point Party Identification and Gallup 3-Point Scale

	Michigan Identification											
Gallup 3-Point Scale	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	Total
Democrat	264	254	21	26	5	0	0	5	2	12	14	603
Independent	7	33	2	173	124	95	3	28	4	10	18	497
Republican	2	8	1	1	3	16	13	151	114	2	2	313
Other (volunteered)	0	1	0	1	4	0	0	0	0	23	3	32
Don't know/refused	1	1	1	0	4	1	0	1	0	5	30	44
Total	274	297	25	201	140	112	16	185	120	52	67	1,489

That responses to the different party identification queries vary could of course be due to the unreliability of the auxiliary measures, not the traditional party identification scale. This possibility, however, is ruled out by the pattern of correlations among the measures of party identification (see tables 9–11). Notice that all but one of the interitem correlations (Pearson *r*) for the NORC data fall between .85 and .89, the lone outlier being a .92 correlation between the 3-point Gallup measure and "basic self-regard" (table 11).[5] Of particular interest is the fact that the correlations between the Michigan scale and the other four items only vary between .87 and .89. In other words, casual inspection of the correlation matrices seems to suggest that the reliabilities are roughly the same for each of the measures. We will return momentarily to a more rigorous assessment of reliability using these patterns of interitem relationships.

Another pattern of note in the cross-tabulation between the Michigan scale and alternative measures is the differential rate of dispersion across the seven partisanship categories. Returning momentarily to table 7, consider the cross-tabulation of the 3-point basic self-regard measure with the Michigan scale. Leaving aside "don't know/other" responses, we find that five of the 160 weak partisans switch parties, as opposed to zero of the 96 leaning independents. None of the pure independents budge at all. The same pattern emerges in the cross-tabulation between the Michigan item and the 7-point self-placement scale (table 5). Again, weak partisans appear more heterogeneous than pure independents, with leaning partisans falling somewhere in between.

Nevertheless, it should be noted that the cross-tabulations presented in tables 5–8 in every case show the party identification scale to be ordinal in nature. None of the five criterion measures for the three NORC subsamples ever produce an "intransitivity" of the sort detected by Petrocik (1974) and documented more recently by Keith et al. (1992). Indeed, in each instance leaning partisans are quite different from weak partisans in terms of the ways in which they respond to alternative measures of partisanship, as are strong and weak partisans. That the ordered series of partisan categories in the Michigan scale do not always produce a monotonic relationship with presidential vote

---

[5] Apparently, when respondents are asked to consider their party identification apart from their past voting behavior (see wording to the basic self-regard item in the Appendix), they tend to focus on the aspect of their partisanship that is akin to party affiliation in politics-as-of-today. Perhaps this explains why political evaluations, such as presidential approval, do not always correlate differently with the Gallup and Michigan measures asked at the same point in time (cf. Abramson and Ostrom [1993] and Bishop, Tuchfarber, and Smith [1992] for opposing findings on this matter). While politics-as-of-today picks up short-run considerations and is apparently more unstable over time (Abramson and Ostrom 1993), it also may shed considerations that were part of past voting decisions, factors that may be correlated with contemporary political evaluations.

Table 9. Means, Standard Deviations, and Correlation Matrix for Experimental Group 1 ($N = 438$)

Variable	Mean	Standard Deviation	Michigan Party ID	7-Point Self-Placement	Gallup 3-Point Scale
Michigan 7-point party ID	3.525	2.052	1.000		
7-point self-placement scale	3.527	1.674	.886	1.000	
Gallup 3-point scale	2.676	1.529	.880	.850	1.000

NOTE.—Data were deleted in listwise fashion in computing the correlations.

Table 10. Means, Standard Deviations, and Correlation Matrix for Experimental Group 2 ($N = 415$)

Variable	Mean	Standard Deviation	Michigan Party ID	Forced-Pair Comparison	Gallup 3-Point Scale
Michigan 7-point party ID	3.398	1.979	1.000		
Forced-pair comparison	2.470	1.705	.874	1.000	
Gallup 3-point scale	2.590	1.590	.873	.857	1.000

NOTE.—Data were deleted in listwise fashion in computing the correlations.

Table 11. Means, Standard Deviations, and Correlation Matrix for Experimental Group 3 ($N = 422$)

Variable	Mean	Standard Deviation	Michigan Party ID	Basic Self-Regard	Gallup 3-Point Scale
Michigan 7-point party ID	3.256	2.012	1.000		
Basic self-regard	2.483	1.580	.871	1.000	
Gallup 3-point scale	2.493	1.571	.878	.924	1.00

NOTE.—Data were deleted in listwise fashion in computing the correlations.

preference may say less about defects of the partisanship measure than about the deficiencies of focusing on the simple bivariate relationship between vote preference and party identification.

The cross-tabulation between the 7-point Gallup item and the party difference scale (table 12) reveals both the transitive ordering of party identification categories and the heterogeneity within each category. Again, weak partisans and leaning independents differ substantially in terms of their locations on the party difference score, although considerable dispersion is evident in both categories. One defect of the party difference scale, as suggested earlier, seems to be the encouragement of indifferent responses. While this accurately gauges the sentiments of those who rate as pure independents by the Gallup measure (despite the fact that the party difference questions make no mention of "independent" identification), it also attracts many respondents in other partisan categories, generating a lumpy distribution in several of the columns of table 12. Nonetheless, the bivariate correlation between the two partisanship measures is .873.

CORRELATIONS WITH DEMOGRAPHIC VARIABLES AND CANDIDATE PREFERENCES

If the various measures of party identification tap the same underlying trait with random error, we should expect to find that they are similarly correlated with determinants and consequences of party affiliation. The only systematic difference we would expect to observe between party identification measures is that less reliable items have more attenuated correlations with assorted demographic variables and vote preferences.[6]

The pattern of correlations is reported in tables 13 and 14. On the whole, the Michigan item is not very different from the other party identification measures in terms of its relationship to the seven criterion variables. The self-placement item is on par with the Michigan measure; the other three items in the NORC study rate slightly lower than the Michigan item in terms of reliability, although the comparison

---

6. The demographic variables are years of education, dummy variables for female, nonwhite, Catholic, Southern residence, and age. (In the statistical analysis that follows, age is entered as a series of three dummy variables.) For the NORC survey, the dependent variable is the sum of responses to five hypothetical 1976 presidential election match-ups. Each response is coded "−1" if the respondent would vote for the Democrat, "1" if he or she would vote for the Republican, and "0" if he or she cannot decide. The five trial heats are Humphrey-Percy, McGovern-Goldwater, Kennedy-Percy, Humphrey-Reagan, and Kennedy-Reagan. The Times-Mirror survey contains only one vote preference question, regarding the intended vote for U.S. House of Representatives in the upcoming Congressional election. This variable is scored "−1" if the respondent intends to vote for the Democratic candidate, "1" if Republican.

between Michigan and Gallup items may be confounded by the long-term versus short-term referent of the questions. The Times-Mirror data provide only a 7-point Gallup measure as a yardstick for comparison, but by this standard it would appear that the 19-point party difference measure is a shade worse in terms of reliability.

Assessing the Measurement Properties of the Michigan Scale

Having examined the relationships between the Michigan party identification measure and alternative measures of party affiliation, we now take up the task of assessing the reliability of the Michigan measure using the NORC data.[7] In order to gauge the degree of measurement error in the Michigan item, we constructed a structural equations model with latent variables (Bollen 1989: Hayduk 1987; Jöreskog and Sörbom 1988) that integrates three sets of cross-sectional information: (i) redundant measures of partisanship, (ii) demographic correlates of partisanship, such as age cohort, race, religion, education, and region, and (iii) presidential candidate preferences across a series of trial heats. Our structural model assumes that partisanship and various demographic variables are mutually intercorrelated causes of candidate preferences. In this respect, our analysis mimics a simple regression model in which candidate preferences are regressed on a list of predictors that includes demographic variables and partisanship (cf. Shanks and Miller 1990). The main departure from standard regression analysis is that partisanship is measured twice in each of the three experimental groups.[8]

Redundant measures of party identification within each experimental group enable us to identify the key parameters of interest and thereby calculate the reliability with which party identification is measured. The availability of three experimental groups, all of which confronted a common set of demographic, party identification, and candidate preference questions, improves the precision with which these parameters may be estimated. Using a multiple-group analysis (Alwin and Jackson 1981; Hayduk 1987; Lomax 1983), we estimated the variance of the

---

[7] Since the Times-Mirror study does not contain this item, we will focus our attention on the NORC data. In a separate analysis not reported here, we analyzed the reliabilities of the two-party identification measures available in the Times-Mirror study. These results, which are available from us on request, show these items to be less reliable than the alternative measures developed by NORC.

[8] Note that our analysis uses only the long-term partisanship questions. To have added the Gallup questions would have risked confounding measurement error with potentially meaningful differences between long-term and short-term partisan self-regard.

Table 12. Cross-Tabulation of Party Difference Score and Gallup 7-Point Scale

	Gallup 7-Point Scale								
Party Difference Score	Strong Democrat	Weak Democrat	Lean Democrat	Independent	Lean Republican	Weak Republican	Strong Republican	Refused or No Answer	Total
−9	315	191	34	5	1	1	2	6	555
−8	44	29	7	3	0	0	0	0	83
−7	47	41	12	2	0	0	0	1	103
−6	25	54	18	3	0	1	0	0	101
−5	19	39	33	1	0	1	0	3	96
−4	18	49	45	4	3	6	1	1	127
−3	6	26	20	4	2	0	0	2	60
−2	10	27	47	9	3	3	2	0	101
−1	7	15	29	11	4	4	2	2	74
0	6	49	83	202	118	38	6	22	524
1	2	4	2	10	38	23	7	2	88

2	1	4	1	4	46	38	4	0	98
3	1	3	1	6	33	31	11	0	86
4	1	1	0	2	37	55	15	1	112
5	0	0	0	1	18	43	25	2	89
6	1	1	1	2	11	49	20	0	85
7	0	0	0	0	8	42	47	0	97
8	0	0	0	0	6	22	31	0	59
9	2	0	3	2	13	76	181	2	279
Don't know (Republican)	2	6	2	2	1	1	2	0	16
Don't know (Democrat)	3	4	9	56	6	4	2	2	86
Refused or no answer	11	13	9	9	10	19	10	3	85
Total	521	556	356	338	359	457	368	49	3,004

NOTE.—See Appendix for scoring.

Table 13. Correlations of NORC Party Identification Measures with Criterion Variables

	Group 1 (*N* = 361)		Group 2 (*N* = 351)		Group 3 (*N* = 345)		Total (*N* = 1,086)	
Criterion Variable	Michigan Party ID	Self-Placement	Michigan Party ID	Forced Pair	Michigan Party ID	Basic Self-Regard	Michigan Party ID	Gallup 3-Point
Vote index	.672	.715	.696	.657	.668	.643	.670	.653
Nonwhite	−.294	−.321	−.290	−.292	−.212	−.237	−.253	−.272
Education	.160	.212	.095	.073	.192	.200	.149	.152
South	.025	.031	.016	.015	−.002	−.005	.015	.001
Catholic	−.167	−.139	−.175	−.134	−.216	−.170	−.187	−.153
Female	.007	.027	−.048	−.038	.081	.053	.015	.010
Age	−.007	−.056	.115	.107	−.029	−.032	.033	.002

Table 14. Correlations of Times-Mirror Party Identification Measures with Criterion Variables ($N = 2{,}073$)

Criterion Variable	Gallup 7-Point	Party Difference Score
Vote intention	.834	.803
Nonwhite	−.288	−.292
Education	.145	.131
South	−.048	−.032
Catholic	−.024	−.013
Female	−.072	−.079
Age	−.038	−.028

NOTE.—Data were deleted in a listwise fashion in computing the correlations.

latent partisan factor, its effect on the vote, and the measurement error variance associated with each of the party identification measures. The model's critical assumption is that the measurement errors associated with the party identification items are random, which is to say that error in each measure of party identification bears no systematic relationship to the other variables in the model or the measurement errors associated with them.[9] The precise specifications of our structural equations with latent variables model are described in a technical appendix available from us.[10]

Since the redundancy with which partisanship is measured produces 8 *df* in each of the three experimental groups and the constraints imposed across groups produces an additional 8 *df*,[11] the entire model has a total of 32 *df*. The overidentification of the model has two benefits. First, it affords us the opportunity to assess the fit between model and data by means of a χ^2 goodness-of-fit test. Second, the large stock of surplus information contributes to the robustness of the estimates in the face of sampling error and certain forms of model misspecification (Green, Goldman, and Salovey 1993). Put simply, the correlations among the partisanship measures, the demographic variables, and candidate preferences enable us to triangulate on the measurement error variance associated with the different measures of party affiliation. To the extent that these correlations reveal that different partisanship measures are associated differently with other variables in the analysis, the statistical fit of the model will be poor, and we will be forced to question whether the assumptions underlying the analysis are valid.

---

[9] The success of the model does not hinge on the assumption that the demographic variables or candidate preference measure are error free. The reliability of partisanship will be estimated consistently provided these variables are measured with random error.

[10] Since the partisanship measures are not continuous and are arguably ordinal rather than metric, we experimented with ordinal factor analysis (Jöreskog and Sörbom 1988). Reestimating the models with weighted least-squares using polychoric correlation matrices as input produce results similar to the maximum likelihood estimates we report (see Green and Palmquist [in press] for a similar finding). For the sake of simplicity, therefore, we rely on more standard techniques for the analysis of covariance structures (Bollen 1989; Long 1983). All of the results reported here are obtained using maximum likelihood estimation.

[11] The variance of party identification, the measurement error associated with the Michigan scale, and the structural disturbance term of the candidate preference equation are constrained to be equal across groups. A similar across-group constraint is imposed on the unstandardized effect of party identification on candidate preference, for a total of four constrained parameters. Since these four parameters are estimated once for three experimental subgroups, there is a savings of 8 *df*.

The results of our analysis, presented in table 15, suggest not only that the fit between model and data is very good (χ^2 [32] = 30.2, p = .56) but also that the reliability of party identification is almost exactly what has been obtained through test-retest methods. The error variance associated with the Michigan party identification scale is .49 (SE = .06), which implies a reliability of approximately .89. By comparison, Green and Palmquist's (in press) analysis of eight multiwave panel studies produced a median error variance of .52 and a median reliability of .87. That the two methodological approaches produce similar findings—a close fit between theory and data and reliability estimates indicating that slightly more than 10 percent of the observed variance in the Michigan scale is random noise—does not prove that these models are correct in their fundamental assumptions. At the same time, the fact that two analyses that rest on different statistical premises produce consistent results means that those who have been skeptical of earlier reliability results (Franklin 1992) can only sustain their skepticism by adopting a significantly more elaborate set of objections to the way in which measurement error is modeled.

As far as the alternative measures are concerned, the only measure whose reliability is at least as great as that of the Michigan scale is the self-placement item. In fact, this item slightly outperformed the Michigan scale in terms of reliability (.92 versus .89), although the difference falls short of statistical significance. None of the other alternative measures of party identification had a reliability exceeding .87.

Selecting Companion Measures to Improve Prediction

In addition to facilitating reliability assessment and parameter estimation in structural equation models, the inclusion of redundant measures

Table 15. LISREL Parameter Estimates for NORC Model

	Unstandardized Factor Loadings	Error Variance	Reliabilities			Estimate	R^2
			Group 1	Group 2	Group 3		
Michigan 7-point party ID	1.000	.486 (.060)	.887	.880	.887	. . .	. . .
7-point self-placement scale	.837	.247 (.051)	.915			. . .	. . .
Forced-pair comparison	.850	.388 (.061)		.870		. . .	. . .
Basic self-regard	.760	.422 (.057)			.838	. . .	. . .
Additional parameters of interest:							
Party identification trait variance	. . .	. . .				3.733	. . .
Unstandardized effect of partisanship	. . .	. . .				1.149 (.044)	. . .
Standardized effect of partisanship	. . .	. . .				.651	. . .
Experimental group:							
Group 1 (self-placement)	. . .	. . .				. . .	58.4
Group 2 (forced-pair comparison)	. . .	. . .				. . .	60.6
Group 3 (basic self-regard)	. . .	. . .				. . .	56.1

NOTE.—Dependent variable is an index of preferences concerning five hypothetical 1976 presidential election match-ups. Fit of the model: χ^2 (32) = 30.17 (p = .559). Numbers in parentheses are standard errors.

of partisanship offers an opportunity to improve the accuracy with which party identification predicts vote preferences. To the extent that party identification is measured with random error, the bivariate relationship between it and the vote will be attenuated. One way to recapture some of the lost predictive capacity of the Michigan party identification measure is to add one or more supplementary measures of partisanship to the regression equation.[12]

One inference to be drawn from the cross-tabulations presented earlier in tables 5–8 is that the most useful supplementary measures are those that can clarify the heterogeneity within the categories of "weak partisans" and discern gradations in partisanship even among those classified by the Michigan measure as "strong partisans." There are, however, practical limitations to the sorts of questions that can be asked in the context of a single survey. It seems awkward to ask both parts of the branched Michigan question twice in the same survey.[13] At most, the Michigan measure may be supplemented by the trunk question of the Gallup scale, as in the NORC study (and even here, there is an uncomfortable similarity in question wording and response format). As we saw in tables 5–8, supplementing the Michigan scale with an additional 3-point Gallup measure does little to illuminate the differences among self-professed partisans. The party difference measure employed in the Times-Mirror study is a step in the right direction, in that it does not duplicate the format of the Michigan questions. Even better, because the responses it elicits do not cluster at the end points of the scale, is the 7-point self-placement scale used in the NORC study.

How much variance in candidate preferences do the auxiliary measures predict over and above the standard Michigan scale? In order to give all of the measures their best opportunity to account for variance, we divide each of the variables into *k* dummy variables, where *k* is one less than the number of valid response categories. The baseline regression model uses only the 7-point Michigan scale as a predictor. Here, candidate preferences are regressed on six dummy variables representing the various categories of the Michigan scale. The alternative regression model adds a series of dummy variables representing particular auxiliary measures of partisanship. The two regression models are then compared by means of an *F*-test to determine whether the inclusion of the supplementary measures contributed significantly to the predictive accuracy of the model.

---

[12] In doing so, one must be careful about interpreting the regression estimates. As Achen (1985) points out, the slope coefficients are biased when redundant measures are introduced. For purposes of prediction, the relevant statistics are the proportion of variance explained (R^2) and the standard error of estimate. The latter statistic penalizes the regression model for the number of predictors it uses. See Johnston (1984).

[13] In the 1987 NES Pilot Study, repeated fully-branched Michigan items caused one-fourth of the sample to comment spontaneously to the interviewers that they had just been asked that question.

The NORC data reveal that the forced-pair comparison, the basic self-regard item, and the 3-point Gallup measure provide modest improvement in predictive accuracy over and above what was obtained using the Michigan scale alone. A series of *F*-tests reveal that each of the three alternative measures, when added to the Michigan 7-point scale, contributes significantly ($p < .05$) to the predictive accuracy of the model. Yet in no case did the inclusion of one or more of these three measures improve the R^2 by more than .03. The only NORC measure to make an appreciable and statistically significant contribution to predictive accuracy is the 7-point self-placement measure, which raised the R^2 from .46 to .53. The success of this measure is consistent with our earlier finding suggesting that the self-placement scale is the most reliable of the alternative measures of party identification.

Turning to the Times-Mirror study, we find that the party difference index provides a marginally improved forecast of intended House of Representatives vote. Together with the 7-point Gallup measure, this item generates an R^2 of .76; when the 7-point Gallup item is the sole regressor, the R^2 drops to .74, a statistically significant ($p < .05$) but small change. In one sense, this disappointing result underscores the limitations of the party difference measure noted above. In addition, these results together with those obtained from the NORC study show how hard it is to improve on the predictive power of the Michigan scale, given its reliability of approximately .89. In order to have an appreciable effect on the coefficient of determination, an auxiliary measure of partisanship must be at least as reliable, a criterion met only by the NORC self-placement scale.

Before drawing this exercise to a close, let us consider one additional predictive model that takes into account the specific pattern of responses to alternative measures of party identification. Earlier, we included dummy variables for the Michigan scale and a supplementary measure on the assumption that the two scales contributed additively to the observed measure of candidate preference. We now create dummy variables for each observed combination of responses to the different party identification questions so as to allow for the possibility that different patterns of answers reflect distinct attitudes. This specification, while a bit eccentric, is not without theoretical significance. To find that this model produces a significant improvement in fit would run counter to the basic premise of our earlier reliability analysis, namely, that responses to various party identification items represent linear manifestations of a single underlying party identification factor measured with error. If responses to different questions interact with one another, then part of what our measurement model dubs random error is really substantively meaningful variation.

Again, *F*-tests were employed to compare the predictive accuracy of two models, one with dummy variables for each configuration of responses and another with dummy variables just for main effects. In each instance, this test furnishes no evidence that interaction terms enhance the model's accuracy ($p < .05$). The only potential exception is the Gallup 3-point scale, which yields statistically significant improved prediction when interactions are included. Although this finding may provide some support for those who argue that the Gallup and Michigan items measure different traits, it is important to note that the interaction effect is partially an artifact of sample size. When the regression results are examined for each experimental group separately, one finds no evidence of any interaction effects.[14] The results, in other words, are compatible with the notion that the various measures of party identification, with the possible exception of the Gallup item, tap a single underlying trait with random error.

Discussion

The reliability assessment undertaken here, in keeping with previous results, indicates that although the Michigan scale is highly reliable compared to other attitude measures (Achen 1975; Krosnick 1991; Krosnick and Berent 1993), it is nonetheless fallible. In the NORC sample roughly 89 percent of the observed variance in the Michigan scale may be attributed to variation in partisanship; the remaining 11 percent is some combination of interviewer biases, misread or misunderstood questions, incomplete translation of the partisanship continuum into a discrete 7-point categorization, coding mistakes, and an inability on the part of respondents to gauge their own partisan sentiment.

To say that the reliability of the Michigan scale is roughly .89 does not appear to pose much of a challenge to previous studies that have presumed this measure to have a reliability of 1.00. But in fact the difference between .89 and 1.00 is sufficiently large to undermine many of the "revisionist" claims against the original notion (Campbell et al. 1960) that partisanship is a trait that remains stable over the course of adulthood. The empirical foundation for many revisionist challenges to the original conception of party identification is that a substantial proportion of a national sample changes its party identification score over a span of a few months (Brody and Rothenberg 1988) or years (Brody 1977) or that the test-retest correlation differs from unity (Allsop and Weisberg 1988; Dreyer 1973). This observation, however, is misleading insofar as it confounds real change in partisanship with change due to response error. When party identification data are collected only minutes apart, thus ruling out the possibility of real change, the same pattern of apparent instability emerges. Our cross-sectional results in tandem with the test-retest results presented by Green and Palmquist (in press) suggest that mismeasurement alone is sufficient to account for much of what passes for real change in partisanship.

---

[14] The Times-Mirror study reveals a small but statistically significant gain from the inclusion of interaction terms between the Gallup measure and the Party Difference Score. This finding, like the NORC results, may be attributable to sample size.

One implication of our analysis is that measurement error must be made an explicit part of models involving partisanship. To do this effectively requires some redundancy in the way partisanship is measured, whether in the form of multiple indicators at a given point in time or repeated measurement across successive waves of a panel study or, ideally, some combination of the two. The present study provides some insights into how alternative measures of party identification might be drafted. While several of the items considered here seemed to meet the criteria of face validity laid out by Converse and Pierce (1985) and Campbell et al. (1960), only the 7-point self-placement measure proves to be relatively immune to "don't know" responses, easily administered in face-to-face interviews, capable of producing a unimodal distribution of responses, and as reliable as the Michigan 7-point scale. Moreover, unlike the other alternative measures, the self-placement scale discerns gradations of partisanship that exist within the "strong" and "weak" categories of the Michigan partisanship classification. While one data set alone cannot firmly establish the attractiveness of a measure, our results suggest that this item deserves a closer look. Given the pivotal role partisanship plays in analyses using data from the American National Election Study and General Social Survey, these face-to-face surveys would do well to supplement and illuminate the traditional Michigan party identification scale.

Appendix

Question Wording

NORC SURVEY

Seven-point Michigan scale. "Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?" If Republican or Democrat: "Would you call yourself a strong [PARTY CHOSEN] or a not very strong [PARTY CHOSEN]?" If Independent: "Do you think of yourself as closer to the Republican or to the Democratic party?"

Seven-point self-placement scale. "On this card is a scale with strong Democrats on one end and strong Republicans on the other, and with independents in the middle. Where would you place yourself on this scale?" The 7-point response scale was marked with the following labels at each scale point: (1) I strongly prefer the Democrats, (2) I prefer the Democrats, (3) I am basically independent but lean toward the Democrats, (4) I am independent, I have no preference for either party, (5) I am basically independent but lean toward the Republicans, (6) I prefer the Republicans, and (7) I strongly prefer the Republicans.

Forced-pair comparison item. "Generally speaking, do you usually think of yourself as a Democrat or a Republican?" Independent was a possible option if volunteered by the respondent.

Basic self-regard item. "No matter how you voted in the last couple of national elections or how you think you might vote in next November's national election—do you basically think of yourself as a Republican, a Democrat, an Independent, or what?"

Three-point Gallup item. "In politics, *as of today* do you consider yourself a Republican, Democrat, or Independent?"

Although our cross-tabulations include all respondents, the correlations presented exclude respondents who answer "other" or "don't know." In addition, those respondents who could not decide if they were strong or not very strong partisans in the Michigan question were excluded. The Michigan question and the self-placement were scored as 7-point scales, ranging from strong Democrats to strong Republicans, with independents who do not lean toward either party in the middle. The other questions were scored as 5-point scales with Democrats scored as 1, independents scored as 3, and Republicans scored as 5.

TIMES-MIRROR SURVEY

Seven-point Gallup scale. "In politics, as of today, do you consider yourself a *Republican*, a *Democrat*, an *Independent*, or what?" If Democrat or Republican: "Would you call yourself a strong [PARTY CHOSEN] or a not very strong [PARTY CHOSEN]?" If "independent," "no preference," "other," or "don't know": "Would you say you lean more to the Republican party or more to the Democratic party?"

Party difference score. "Now I'd like to ask you a question about how you regard yourself. On a scale from 1 to 10, where '10' represents a description that is perfect for you, and '1' a description that is totally wrong for you, how well do each of the following describe you? . . . A Republican. . . . A Democrat." Each respondent's party difference score is computed by subtracting the score given to the Democrats from the score given to the Republicans. Respondents who declined to provide a score for either party are excluded from the statistical analysis.

References

Abramson, Paul R., and Charles W. Ostrom, Jr. 1991. "Macropartisanship: An Empirical Reassessment." *American Political Science Review* 85:181–92.

———. 1992. "Question Wording and Macropartisanship: Response." *American Political Science Review* 86:481–86.

———. 1993. "Question Wording and Partisanship: Change and Continuity in Party Loyalties during the 1992 Election Campaign." Paper presented at the annual meeting of the American Political Science Association, Washington, DC, September 2–5.

Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69:1218–31.

———. 1985. "Proxy Variables and Incorrect Signs on Regression Coefficients." *Political Methodology* 11:299–316.

Allsop, Dee, and Herbert F. Weisberg. 1988. "Measuring Change in Party Identification in an Election Campaign." *American Journal of Political Science* 32:996–1017

Alwin, Duane F., and David J. Jackson. 1981. "Applications of Simultaneous Factor Analysis to Issues of Factorial Invariance." In *Factor Analysis and Measurement in Sociological Research: A Multi-dimensional Perspective*, ed. David J. Jackson and Edgar F. Borgatta, pp. 249–79. Beverly Hills, CA: Sage.

Asher, Herbert B. 1974. "Some Consequences of Measurement Error in Survey Data." *American Journal of Political Science* 18:469–85.

Barnes, Samuel H., M. Kent Jennings, Ronald Inglehart, and Barbara Farah. 1988. "Party Identification and Party Closeness in Comparative Perspective." *Political Behavior* 10:215–231.

Bishop, George F., Alfred J. Tuchfarber, and Andrew E. Smith. 1992. "Question Wording and Partisanship: An Experimental Test of the Hypothesis." Paper presented at the annual meeting of the American Political Science Association, Chicago, September 3–6.

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley & Sons.

Brody, Richard A. 1977. "Stability and Change in Party Identification: Presidential to Off-years." Paper presented at the annual meeting of the American Political Science Association, Washington, DC, September 1–4.

Brody, Richard A., and Lawrence S. Rothenberg. 1988. "The Instability of Partisanship: An Analysis of the 1980 Presidential Election." *British Journal of Political Science* 18:445–65.

Budge, Ian, Ivor Crewe, and Dennis Farlie, eds. 1976. *Party Identification and Beyond: Representations of Voting and Party Competition*. London and New York: Wiley.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Campbell, James E., Mary Munro, John R. Alford, and Bruce A. Campbell. 1986. "Partisanship and Voting." In *Research in Micropolitics*, ed. Samuel Long, 1:99–126. Greenwich, CT: JAI Press.

Converse, Philip E., and Roy Pierce. 1985. "Measuring Partisanship." *Political Methodology* 11:143–66.

Dennis, Jack. 1988. "Political Independence in America. I. On Being an Independent Partisan Supporter." *British Journal of Political Science* 18:77–109.

Dreyer, Edward C. 1973. "Change and Stability in Party Identifications." *Journal of Politics* 35:712–22.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Franklin, Charles H. 1992. "Measurement and the Dynamics of Party Identification." *Political Behavior* 14:297–309.

Franklin, Charles H., and John E. Jackson. 1983. "The Dynamics of Party Identification." *American Political Science Review* 77:957–73.

Green, Donald Philip. 1990. "The Effects of Measurement Error on Two-Stage Least-Squares Estimates." *Political Analysis* 2:57–74.

Green, Donald Philip, Susan Lee Goldman, and Peter Salovey. 1993. "Measurement Error Masks Bipolarity of Affect Ratings." *Journal of Personality and Social Psychology* 64:1029–41.

Green, Donald Philips, and Bradley Palmquist. 1990. "Of Artifacts and Partisan Instability." *American Journal of Political Science* 34:872–902.

———. In press. "How Stable Is Party Identification?" *Political Behavior.*

Hayduk, Leslie A. 1987. *Structural Equation Modeling with LISREL.* Baltimore and London: John Hopkins University Press.

Heise, David R. 1969. "Separating Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34:93–101.

Jackson, John E. 1975. "Issues, Party Choices, and Presidential Votes." *American Journal of Political Science* 19:161–85.

Johnston, J. 1984. *Econometric Methods.* 3d ed. New York: McGraw-Hill.

Jöreskog, Karl G., and Dag Sörbom. 1988. *LISREL VII.* Mooresville, IN: Scientific Software.

Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark C. Westley, and Raymond E. Wolfinger. 1992. *The Myth of the Independent Voter.* Berkeley and Los Angeles: University of California Press.

Kenney, Patrick, and Tom Rice. 1988. "The Evaporating Independents." *Public Opinion Quarterly* 52:231–93.

Krosnick, Jon A. 1991. "The Stability of Political Preferences: Comparisons of Symbolic and Nonsymbolic Attitudes." *American Journal of Political Science* 35:547–76.

Krosnick, Jon A., and Matthew K. Berent. 1993. "Comparisons of Party Identification and Policy Preferences: The Impact of Survey Question Format." *American Journal of Political Science* 37:941–64.

Lomax, Richard G. 1983. "A Guide to Multiple-Sample Structural Equation Modeling." *Behavior Research Methods and Instrumentation* 15:580–84.

Long, J. Scott. 1983. *Confirmatory Factor Analysis.* Beverly Hills, CA: Sage.

MacKuen, Michael B., Robert S. Erickson, and James A. Stimson. 1989. "Macropartisanship." *American Political Science Review* 83:1125–42.

———. 1992. "Question Wording and Macropartisanship." *American Political Science Review* 86:475–81.

Markus, Gregory B., and Philip E. Converse. 1979. "A Dynamic Simultaneous Equation Model of Electoral Choice." *American Political Science Review* 73:1055–70.

Miller, Warren E. 1991. "Party Identification, Realignment, and Party Voting: Back to the Basics." *American Political Science Review* 85:557–68.

Nie, Norman H., Sidney Verba, and John R. Petrocik. 1979. *The Changing American Voter,* expanded ed. Cambridge, MA: Harvard University Press.

Niemi, Richard G., David R. Reed, and Herbert F. Weisberg. 1991. "Partisan Commitment: A Research Note." *Political Behavior* 13:213–21.

Page, Benjamin I., and Calvin C. Jones. 1979. "Reciprocal Effects of Policy Preferences, Party Loyalties, and the Vote." *American Political Science Review* 73:1071–89.

Palmquist, Bradley, and Donald P. Green. 1992. "Estimation of Models with Correlated Measurement Errors from Panel Data." *Sociological Methodology 1992,* ed. Peter V. Marsden, pp. 119–46. Washington, DC: American Sociological Association.

Petrocik, John R. 1974. "An Analysis of Intransitivities in the Index of Party Identification." *Political Methodology* 1(3):31–47.

Schickler, Eric, and Donald Philip Green. 1993. "Issues and the Dynamics of Party Identification: A Methodological Critique." Revised version of paper presented at the annual meeting of the Midwest Political Science Association, Chicago, April 16–19.

Shanks, J. Merrill, and Warren E. Miller. 1990. "Policy Direction and Performance Evaluation: Complementary Explanations of the Reagan Elections." *British Journal of Political Science* 20:143–235.

Sullivan, John L., and Stanley Feldman. 1979. *Multiple Indicators: An Introduction.* Beverly Hills, CA: Sage.

Wiley, David E., and James A. Wiley. 1970. "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35:112–17.

Wiley, James A., and Mary Glenn Wiley. 1974. "A Note on Correlated Errors in Repeated Measurements." *Sociological Methods and Research* 3:172–88.

SURVEY EVIDENCE ON EXCESSIVE PUBLIC PESSIMISM ABOUT THE FUTURE BEHAVIOR OF UNEMPLOYMENT

PAMI DUA
DAVID J. SMYTH

Abstract Two of the questions in the Surveys of Consumer Attitudes undertaken by the Survey Research Center of the University of Michigan deal with households' expectations about inflation and the change in unemployment. We compare quarterly time series of the mean responses with the actual behavior of inflation and unemployment to see whether households are overly optimistic or pessimistic about the future behavior of inflation and unemployment. We find an asymmetry. Over the 21 years of our analysis, on average the public is neither overly optimistic nor overly pessimistic about future inflation. However, the public has been significantly too pessimistic about the future behavior of unemployment. These results have an important implication for macroeconomic policy. If politicians respond to complaints from the public about inflation and unemployment they will target policy instruments at unemployment to a greater extent than if the public was not so pessimistic about the behavior of unemployment. The result may be a higher rate of inflation than if the public were not mistakenly pessimistic about the behavior of unemployment.

I. Introduction

This article presents evidence showing that the public is overly pessimistic about the future behavior of unemployment, but not about the future behavior of inflation. This may cause macroeconomic policy to

---

PAMI DUA is associate professor of economics at the University of Connecticut. DAVID J. SMYTH is professor of economics at Louisiana State University. The authors would like to thank three anonymous referees for comments on three earlier versions of this article. They also benefited from comments by Roy Batchelor. Smyth's research was supported by the LSU Foundation.

Public Opinion Quarterly Volume 57:566–574 © 1993 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/93/5704-0008$02.50

be biased toward measures that try to reduce unemployment at the expense of inflation.

We use data on households' expectations of the rate of inflation and changes in unemployment over the next 12 months and the actual behavior of inflation and unemployment. The data on expectations are taken from the responses of households to the Surveys of Consumer Attitudes conducted by the Survey Research Center (SRC), University of Michigan. The surveys were run in the mid-month of each quarter until 1977 and monthly since then. Since continuous data for the mid-month of quarters are available only from 1969, our analysis covers the forecasts made from the first quarter of 1969 to the first quarter of 1990. These are 4-quarter-ahead predictions for the first quarter of 1970 to the first quarter of 1991. We thus have 85 quarterly observations. Over this time period, the sample sizes have ranged from 500 to 2,482 per survey, with sample sizes declining over time.

II. Expected and Actual Inflation

The data on the expected inflation rate are from the SRC's tabulations of the mean responses to questions on expected inflation. From 1966 to 1977, the questions were, "During the next 12 months do you think that prices in general will go up, or go down, or stay the same?" and "How large a price increase do you expect? Of course nobody can know for sure, but would you say that a year from now prices will be about 1% or 2% higher, or 5%, or closer to 10%, or what?" Since 1978, the first question is the same as that asked before. However, in the follow-up question, the respondents are asked to state a quantitative change: "By about what percent do you expect prices to go up, on the average, during the next 12 months?" Through 1977 we use the series quantified by Juster and Comment (1980). Thereafter we use the mean responses reported by the SRC.

The expected inflation rate is compared with the actual inflation rate over the past 12 months computed from the Consumer Price Index (all items, all urban consumers, 1982–84 = 100). Data on the price index are taken from *Business Statistics*, 1961–88, and recent issues of the *Survey of Current Business*.

Figure 1 graphs the actual rate of inflation and the expected inflation rates made 12 months earlier. The expected inflation series does not forecast actual inflation perfectly, but the two series do move together, the correlation coefficient between the two being 0.783.[1]

---

[1] The root mean square error for the inflation forecasts is 1.866. For discussions of the characteristics of the SRC inflation forecasts see Batchelor and Dua (1989), Rich (1989), Smyth (1992b), and the references cited therein.

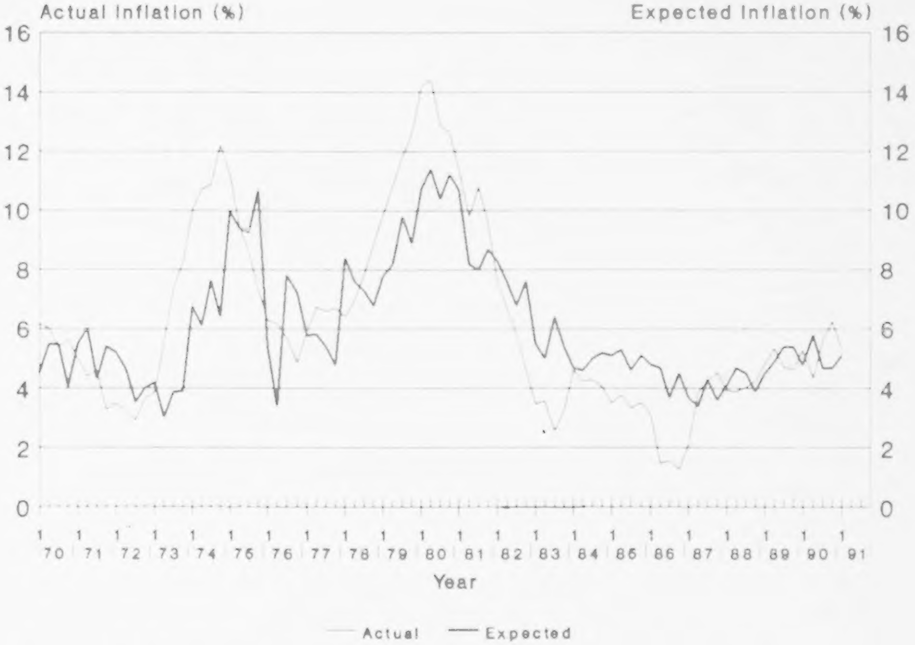


Figure 1. Expected and actual inflation, first quarter of 1970 through the first quarter of 1991.

Table 1 is a two-by-two contingency table for the change in the inflation rate. Data for each quarter are classified in two ways—by whether actual inflation rose compared with 12 months ago and by whether respondents correctly or incorrectly forecast the direction of change. For the 44 quarters during which inflation rose, households expected a rise in 31 quarters. The picture is quite similar for the quarters during which inflation fell. For the 41 quarters during which inflation fell, households correctly predicted a decrease 26 times.

Table 1. Contingency Table between the Expected and Actual Change in Inflation

Forecast of Direction of Change	Actual Change in Inflation		
	≤0	>0	Total
Incorrect	15	13	28
Correct	26	31	57
Total	41	44	85

We use a chi-square test to test the null hypothesis that the proportion of forecasts for which the direction of change proves incorrect is the same when actual inflation is increasing as when it is decreasing.[2] A test statistic of 0.476 with 1 df is obtained, so the null hypothesis is not rejected at the 5 percent significance level. Households are equally accurate at predicting the direction of change of inflation when inflation is falling as when it is rising.

III. Expected and Actual Changes in the Unemployment Rate

The question on the expected unemployment rate in the Surveys of Consumer Attitudes is, "How about people out of work during the coming 12 months—do you think that there will be more unemployment than now, about the same, or less?" The index computed to denote the direction of expected change is

$$\text{Index} = 100[(\text{more} + 0.5\text{same})/(\text{more} + \text{same} + \text{less})], \qquad (1)$$

where "more" is the percentage of households responding "more unemployment," "same" is the percentage responding "same unemployment," and "less" is the percentage responding "less unemployment."[3] In this index, the category "same" is divided equally between "more" and "less." The index employs raw data and is preferred to a series quantified using, say, the Carlson-Parkin (1975) technique that requires an assumption on the distribution of expectations among respondents and the imposition of unbiasedness. We repeated the analysis described below using three alternative measures, the proportion saying more, the proportion saying less, and the index 100[more/(more + less)]. We obtained similar conclusions to that with the use of the index in equation (1) and so do not report details here. An index of 50 means that as many respondents expect unemployment to rise over the next 12 months as to fall. An index over (under) 50 means that more (less) respondents expect an increase in unemployment than expect a fall.

Since households are asked about the direction of expected change

[2] The particular chi-square test that we use is that described in Berenson, Levine, and Rindskopf (1988, sec. 11.4.1).

[3] An alternative index that is widely used in analysis of public opinion data is the "balance index" given by

$$\text{IndexBal} = 100(\text{more} - \text{less})/(\text{more} + \text{same} + \text{less}).$$

The two indexes are related as follows:

$$\text{IndexBal} = 2(\text{Index} - 50),$$

so Index is simply a linear transformation of the balance index.

in the unemployment rate, the index is compared to the annual change in the unemployment rate expressed in percentage points—that is, if unemployment falls from 7.0 percent to 6.5 percent the change in the unemployment rate measure is −0.5. The data on the actual unemployment rate (all civilian workers, seasonally adjusted) are from *Business Statistics*, 1961–88, and recent issues of the *Survey of Current Business*.

Figure 2 plots the index and the actual change in the unemployment rate. The scales for the two variables have been chosen so that an index of 50 corresponds to no change in the unemployment rate. It is clear from figure 2 that the index is usually above 50. The correlation coefficient between the index and change in unemployment is 0.647. The two series move together but less closely than the inflation series.[4]

Table 2 is a two-by-two contingency table in which data for each quarter are classified in two ways, by whether actual unemployment rose compared with 12 months ago and by whether a majority of respondents correctly or incorrectly forecast the direction of change. For the 34 quarters during which unemployment actually rose, households expected a rise on all but one occasion. The picture is very different for those quarters during which unemployment actually fell. Unemployment fell during 51 quarters, but households expected a decrease only 18 times and predicted an increase 33 times. Thus more than two-thirds of the times when unemployment fell, more households were predicting an increase than were predicting a decrease.

To see whether this asymmetry is significant we compare the pattern in table 2 with that generated by a null hypothesis that the proportion of forecasts of the direction of change that prove incorrect is the same when the actual change in unemployment is negative as when it is positive. The resulting chi-square test statistic is 32.43, so the null hypothesis is easily rejected at the 1 percent significance level. Households are significantly less accurate at predicting the direction of change of unemployment when unemployment is falling than when it is rising.[5]

---

[4] We cannot report a root mean square error for forecast unemployment as we could for inflation.

[5] We repeated our analysis for two further definitions of the actual unemployment rate using data supplied by the Bureau of Labor Statistics (BLS): first, aggregate hours lost by the unemployed and persons on part-time for economic reasons as a percentage of potentially available labor force hours; and, second, the BLS U-7 definition, the sum of total full-time job seekers plus 0.5 times part-time job seekers plus 0.5 times total on part-time for economic reasons as a percentage of the sum of the civilian labor force plus discouraged workers less 0.5 times the part-time labor force. The results obtained were similar to those reported in the text. We also measured the direction of change, regardless of whether the extent of that change was sufficient to move the series above or below 50 on the index. When this procedure was adopted, our conclusions were unchanged.

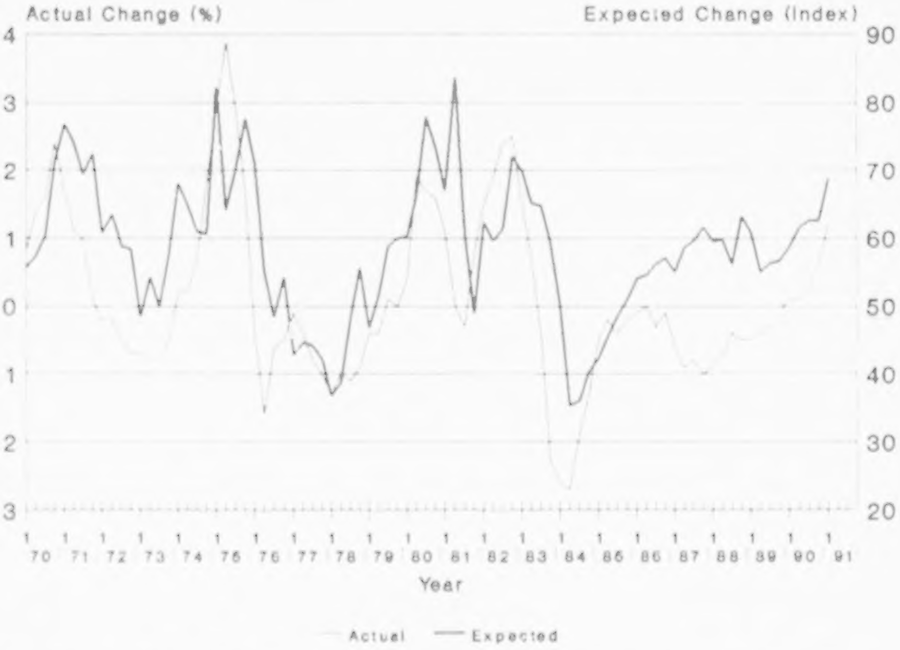


Figure 2. Expected and actual change in unemployment, first quarter of 1970 through the first quarter of 1991.

IV. Implications for Macroeconomic Policy

Our analysis suggests an asymmetry. Over the 21 years of our analysis, on average the public is neither overly optimistic nor overly pessimistic about future inflation. However, the public has been significantly too pessimistic about the future behavior of unemployment. Thus on aver-

Table 2. Contingency Table between the Index of Expected Behavior and the Actual Change in Unemployment

Forecast of Direction of Change	Actual Change in Unemployment		
	≤0	>0	Total
Incorrect	33	1	34
Correct	18	33	51
Total	51	34	85

age the public will complain more about unemployment than it would if it correctly expected future changes in unemployment. If politicians respond to the complaints of the public, they will target policy instruments at unemployment to a greater extent than if the public was not so pessimistic about the behavior of unemployment.

Increased targeting of unemployment relative to inflation leads to short-run reductions in unemployment at the expense of increases in the inflation rate. This will have some short-run benefits in terms of increased output, and some economists may view the resulting effects on income distribution as beneficial. However, there will be long-run costs that we believe will substantially outweigh the short-run benefits when economic capacity is underutilized.

In the long run, the rate of unemployment is constrained by the natural rate of unemployment, and the actual level of unemployment cannot remain below the natural rate of unemployment without ever accelerating inflation. At the natural rate of unemployment the inflation rate depends, inter alia, on policy measures, so policy measures aimed at reducing unemployment end up increasing the rate of inflation. Thus the public's unjustified pessimism about unemployment causes the inflation rate at the natural rate of unemployment to be higher than if the public was not pessimistic about unemployment.

Our analysis provides empirical support for a key assumption in a class of theoretical models that has become popular in recent years. In these models policymakers have an objective function that gives weight to the reduction of unemployment even below the natural rate.[6] The usual justification for the unemployment assumption is the presence of unemployment compensation, distortionary taxes on labor income, and the like. Our explanation for the unemployment assumption is that policymakers are responding to the misconceptions of the public, which is overly pessimistic about the future behavior of unemployment.

Whether there are more social costs to a policy that keeps unemployment below its natural level at the expense of ever-accelerating inflation is a matter of debate among economists. Both sides agree that inflation affects the distribution of income among members of the population. More controversial is whether inflation has indirect effects on the level of output and its rate of growth. Inflation distorts price signals and affects the ability of agents in both the private and public sectors to plan efficiently. Participants in the economic process waste time and resources trying to cope with inflation. Recent research finds these costs to be substantial.[7] Consequently we believe that the public's excessive pessimism about the future behavior of unemployment may have a significant adverse effect on the economy.

---

[6] Barro and Gordon (1983) provide the starting point for most recent work. For a simple exposition of a basic model see Cukierman (1992, chap. 3).

There are three elements of the hypothesis in this article. First, there is an asymmetry in households' perceptions of unemployment relative to inflation, with households being overly pessimistic about unemployment. Second, economic policymakers act on these overly pessimistic unemployment expectations. Third, this results in a lower unemployment rate and a higher inflation rate than otherwise would be the case, which may be detrimental to the economy.

To adequately assess the macroeconomic policy implications of our results we need to know whether or not unemployment rates below the natural rate have been a frequent occurrence. If the unemployment rate is below the natural rate, any attempt to lower the unemployment rate would raise the inflation rate. Using the data on the natural rate reported in Braun (1990), we find that the unemployment rate is at or below the natural rate in 49 of the 85 quarters examined, that is, more than 50 percent of the time.[8] Moreover, in 43 of these 49 quarters the index computed to denote the direction of expected change in unemployment is greater than 50. This means that in half of the quarters examined the actual unemployment rate is at or below the natural rate *and* more respondents expect unemployment to rise than to fall. Thus if policymakers respond to the public's pessimism by reducing unemployment the result will be increased inflationary pressure half the time.[9]

References

Barro, Robert J., and David B. Gordon. 1983. "A Positive Theory of Monetary Policy in a Natural Rate Model." *Journal of Political Economy* 91:589–610.

---

[7] See Economic Report of the President (1985, p. 46) for a brief discussion and Howitt (1990) for an extended discussion of the costs of inflation. See Smyth (1992a), McMillin and Smyth (1993), and Smyth (1994) for empirical evidence on the harmful effects of inflation on output growth.

[8] The natural rate of unemployment increased from 5.9 percent in 1969 to a peak of 7.4 percent from 1975 to 1979 and then decreased to 6.2 percent in 1990. Braun's estimates end with the first quarter of 1988. We extended his series by assuming that the natural rate decreased by 0.1 percent every year. We believe that the natural rate has drifted lower in recent years because of demographic factors (fewer teenagers, who have high unemployment rates, and more older people) and productivity changes. For a discussion of the demographic and productivity issues, see Braun (1984).

[9] At the time of the most recent revision of this article (June 1993) the latest unemployment rate that is available is 6.9 percent for May 1993 compared with a natural unemployment rate of approximately 6.0 percent. Thus a reduction in the actual rate of unemployment would not cause continually accelerating inflation at the present time.

Batchelor, Roy, and Pami Dua. 1989. "Household versus Economist Forecasts of Inflation: A Reassessment." *Journal of Money Credit and Banking* 21:252–57.

Berenson, Mark L., David M. Levine, and David Rindskopf. 1988. *Applied Statistics: A First Course*. Englewood Cliffs, NJ: Prentice Hall.

Braun, Steven N. 1984. "Productivity and the NIIRU (and Other Phillips Curve Issues)." Working Paper no. 34, National Income Section and Wages, Prices and Productivity Section, Board of Governors of the Federal Reserve System.

———. 1990. "Estimation of Current-Quarter Gross National Product by Pooling Preliminary Labor-Market Data." *Journal of Business and Economic Statistics* 8:293–304.

Carlson, John A., and Michael Parkin. 1975. "Inflation Expectations." *Economica* 42:123–38.

Cukierman, Alex 1992. *Central Bank Strategy, Credibility, and Independence*. Cambridge, MA: MIT Press.

Economic Report of the President. 1985. *Economic Report of the President*. Washington, DC: Government Printing Office.

Howitt, P. 1990. "Zero Inflation as a Long-Term Target for Monetary Policy." In *Zero Inflation: The Goal of Price Stability*, ed. R. G. Lipsey, pp. 67–108. Toronto: C. D. Howe Institute.

Juster, F. Thomas, and Robert Comment. 1980. "A Note on the Measurement of Price Expectations." Discussion Paper, University of Michigan, Institute for Social Research.

McMillin, W. Douglas, and David J. Smyth. 1993. "A Multivariate Time Series Analysis of the Aggregate Production Function." Discussion Paper, Louisiana State University, Department of Economics.

Rich, Robert W. 1989. "Testing the Rationality of Inflation Forecasts from Survey Data: Another Look at the SRC Expected Price Change Data." *Review of Economics and Statistics* 71:682–86.

Smyth, David J. 1992a. "Inflation and the Growth Rate in the United States' Natural Output." *Applied Economics* 24:567–70.

———. 1992b. "Measurement Errors in Survey Forecasts of Expected Inflation and the Rationality of Inflation Expectations." *Journal of Macroeconomics* 14:439–48.

———. 1994. "Inflation and Growth." *Journal of Macroeconomics* (in press).

THE POLLS—A REVIEW

RUSH TO JUDGMENT? FAST REACTION POLLS IN THE ANITA HILL–CLARENCE THOMAS CONTROVERSY

DIANNE RUCINSKI

The nomination and subsequent confirmation of Clarence Thomas to the Supreme Court proved to be one of the more controversial issues of 1991. The controversy was due less to his legal philosophy (although that did generate some debate), or his unwillingness to discuss his legal opinions and beliefs, or his qualifications (Thomas received the lowest recommendation for any Supreme Court Justice from the American Bar Association but was praised by President Bush as the most qualified candidate available) than to allegations that he sexually harrassed Anita Hill.

Thomas supervised Hill when he directed the Equal Employment Opportunity Commission. The accusation surfaced publicly days before the final Senate vote was to occur, causing the Judiciary Committee to recommend delaying the vote for 1 week until the charges could be further explored. The charges were examined during hearings, portions of which were aired live on cable, network TV, and radio to an audience estimated to be as large as 27 million homes. As in many political controversies, public opinion polls were highly prominent, if not crucial, ingredients in the debate leading up to the Senate confirmation vote.

The initial portrayal of public opinion by elites and in media discourse was one of consensus: that a majority of the American people, both men and women, believed Thomas's denials of Hill's charges, that African-Americans overwhelmingly rallied behind Thomas after the charges were leveled, and that a majority of the public supported his confirmation. Thomas's support in the polls was cited by those

---

DIANNE RUCINSKI is assistant professor of Communication Studies at the University of Iowa. She would like to thank those polling organizations who cooperated with her requests for information for this review. Additionally, she would like to thank Robert Groves, Michael Kagay, Jacob Ludwig, and the editors of the polls section for helpful comments. None of them bears responsibility for the analyses and interpretation presented here.

Public Opinion Quarterly Volume 57:575–592 © 1993 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/93/5704-0002$02.50

favoring his confirmation as evidence of his veracity and fitness for the Supreme Court.

And yet the controversy has continued. A little less than 1 year after the confirmation vote, three national polls suggested that opinion had shifted from supporting Thomas's version of the truth to supporting Hill's. Two of the three found that more people now believed Anita Hill and one reported that equal portions believed Hill and Thomas. The apparent shift in support from Thomas to Hill has produced much discussion and speculation.

Did the year-after polls reflect a real change in public perception of Anita Hill's charges? Or was the initial support for Thomas really "soft" or illusory—based on uncrystallized opinions? Did nonresponse bias affect the poll findings, particularly for polls conducted during the hearings when interviewing often was completed in one night? Just how good were the estimates of public support for Thomas and Hill at either point in time? Answers to such questions, though difficult to come by, are important to understanding one of the most intriguing and divisive events in American history.

Review of the Hill-Thomas Polls

Recognizing the centrality of polls not only in the initial discussion of Thomas's confirmation but also in the construction of historical arguments about the Hill-Thomas matter, this article examines aspects of the polls and their documentation with an eye to informing the debate. It also provides a glimpse into the workings of polls conducted under enormous time pressure in a highly charged and fluid environment.

The Hill-Thomas polls offer an interesting case study of a particularly prominent form of survey research, highlight the considerable trade-offs involved in it, and illuminate controversies in the survey profession's norms and practices. The review begins with a chronology of the Hill-Thomas matter. It then offers an exploration of the methods employed in these polls in the context of the survey methodological literature.

Chronology

A mere 9 days elapsed from the time Anita Hill's allegations became public, October 6, 1991, to the final Senate confirmation vote, October 15, 1991. National Public Radio and *Newsday* broke the sexual harassment story, and the controversy dominated news programs during that period. By Tuesday, October 8, the Senate decided to delay the confirmation vote for a week in order to investigate the charges through formal hearings. The hearings began on Friday morning, October 11, with Thomas's opening statement. Hill followed with approximately 6 hours of testimony. She did not appear to testify again. Thomas returned again Friday evening during prime time and continued refuting Hill's charges on Saturday. Four witnesses for Hill appeared on Sunday morning, and 29 Thomas supporters appeared to give testimony for the remainder of the day.

In that 9-day period 11 national polls focused exclusively on the Hill-Thomas issue were conducted, beginning with an ABC/*Washington Post* poll on October 8, 1991, and ending with an NBC/*Wall Street Journal* poll on October 15, 1991. The majority of polls were conducted on the weekend of the hearings into Hill's allegations. Nearly a year later, in mid-September and early October of 1992—in the middle of the presidential election campaign—three national polls included questions on the Hill-Thomas matter, marking the anniversary of the Thomas confirmation.

Inside the Hill-Thomas Polls

This chronology serves as background for an examination of the methods employed in the Hill-Thomas polls. These polls embody the current state of the art in fast reaction polls and, thus, it is instructive to look at their inner workings in relation to the more general survey methodological literature. Such an examination reveals controversial decisions in study execution and unknowns and disagreements within the survey community on professional norms and practices.

DISCLOSURE OF METHODS

We begin the examination with a look at sampling and respondent selection procedures. Perhaps the most striking feature of table 1, which contains sampling information for the Hill-Thomas polls, however, is that a number of prominent polling organizations either could not or would not provide information for this review about the procedures they employed. Numerous requests for information—beginning some 8 months prior to the review being finalized—had to be made before information was released by some organizations. For those that did reply, not all of the requested information was provided in every case.

Why should polling organizations provide methodological information on request, and why do some fail to do so? Professional codes

Table 1. Sampling Information for Hill-Thomas 1991 and 1992 Polls

Organization and Date	Household Selection	Respondent Selection	Reported Response Rate (%)	*N*
CBS/*New York Times*:				
10/9/91	9/91 panel	Panel	68[a]	512
10/13/91	9/91 panel	Panel	66	501
10/14/91	Panel	Panel	66	497
Chilton for ABC/*Washington Post*:				
10/12/91	N.A.	N.A.	N.A.	513
10/13/91	N.A.	N.A.	N.A.	506
Gallup:				
10/10–13/91	RDD	Gender quota with substitution	54[b]	1,003
10/14/91	RDD	Gender quota with substitution	47[b]	778
Los Angeles Times:				
10/12–13/91	RDD	Modified last birthday	52[b]	126
Gordon S. Black Corp. for *USA Today*:				
10/13/91	N.A.	N.A.	N.A.	758
Louis Harris and Associates:				
10/9–13/91	N.A.	N.A.	N.A.	1,254
Hart-Teeter for NBC/*Wall Street Journal*:				
10/15/91	RDD	Modified last birthday	N.A.	832
ICR Survey Research for Princeton Survey Research Associates and *U.S. News and World Report*:				
9/11–15/92	Single-stage stratified RDD	Modified last birthday	27–36[b]	1,002
Hart-Breglio for NBC/*Wall Street Journal*:				
9/12–15/92	RDD	Modified last birthday	N.A.	1,506
Gallup:				
10/1–3/92	RDD	Gender quota with substitution	N.A.	1,243

NOTE.—RDD = random digit dialing.

[a] CBS/NYT reinterviewed respondents originally interviewed on the Thomas nomination September 3–5, 1991. Figures given are recontact rates out of a possible 100 percent. The response rate for the original September sample ranged between 56 and 71 percent, depending on how many "ring, no answer" cases were eligible.

[b] The reported response rates are those provided by each organization. See text for discussion of response rate calculation.

written for various survey research organizations—for example, the American Association for Public Opinion Research (AAPOR), the World Association for Public Opinion Research (WAPOR), the Council of American Survey Research Organizations (CASRO), and the National Council of Public Polls (NCPP)—all contain provisions enjoining members to provide information on how their surveys are conducted if the survey findings are released publicly. The codes vary in coverage and specificity, but their purpose is to further professional competence, integrity, and public trust. Disclosure of methods permits open examination of the processes through which conflicting survey findings may result (in lieu of a uniform code of survey practice) and offers protection against misrepresentation of survey findings. The evolution of disclosure codes is reviewed by Gollin (1992), Hollander (1992), and Turner and Martin (1984).

Disclosure practices at the time of the release of poll information vary considerably, with some in the profession arguing that either more or less information about survey methods ought to be published along with poll findings, depending on assumptions about poll consumers' interest and ability to understand methodological details (see Miller, Merkle, and Wang [1991] for a review). There appears to be no disagreement, however, that methodological information ought to be available to those who request it if the published reports do not contain the necessary background for understanding how a poll's findings were produced.

What constitutes the "necessary background" for understanding a poll's findings? The AAPOR standards for minimal disclosure, upon which requests for information for this review were based, include sponsorship of the poll, exact question wording and context and instructions to interviewers, definition of the population and sampling frame, description of the sample selection procedures, sample size, completion rate, and screening procedures. Such information is relevant to understanding poll findings because each factor has been shown to be a potent agent in shaping them.

Why, then, do some polling organizations not provide methodological information when it is requested? Obviously, one answer is that providing information opens up the possibility of criticism and second-guessing. Those organizations that did provide at least some information for this review—CBS/*New York Times*, Gallup, Louis Harris and Associates, *Los Angeles Times*, NBC/*Wall Street Journal* (polls conducted by Hart-Teeter and Hart-Breglio), Princeton Survey Research Associates—therefore are to be commended for opening up their operations to scrutiny. (Princeton Survey Research Associates was particularly prompt and thorough in responding to my requests, despite the fact that records had to be gathered from ICR Survey Research, which had conducted the field work for its poll.) Their willingness and ability to answer such requests is a significant contribution to survey professionalism. Their work, moreover, should not be viewed more critically just because they were able and willing to provide documentation.

Another factor in the failure to make methodological information available is the fact that doing so involves real costs. Detailed records must be kept and archived over years. People must be assigned to the task of answering inquiries and putting data files together. Oftentimes, the task of providing information falls to individuals who are busy with ongoing surveys, making the work particularly onerous. Establishing a system for documenting survey procedures and providing information when it is requested may come with a significant price tag.

The norms of professional conduct permeate some organizations so much that they are willing to pay the price. In other organizations, the norms are less important or are irrelevant. Those directly involved in conducting polls may see the value in thorough documentation, but upper management or the client may not.

Among those who provided no information for this review, only ABC flatly refused without explanation. (Chilton, a subsidiary of ABC at which the field work for the ABC/*Washington Post* polls is conducted, was willing to answer information requests if granted permission to do so.) The Gordon S. Black Corporation cited costs and offered to provide information for a consulting fee.

Public opinion poll archives such as the one at the Roper Center are a possible solution to the problems of documenting and disseminating methodological information. With a standard, detailed set of requirements for such information, an archive could serve as a clearinghouse for requests that otherwise must be directed to individual survey organizations. Unfortunately, the Roper Center receives information in varying levels of detail—it does not, for example, receive information from contributing survey organizations about sampling, respondent selection procedures, call records, or any special field procedures. These items of information were included among those I sought—based on the AAPOR code—in my attempt to document the Hill-Thomas polls. It appears that such information does not receive the same priority as question wording when archived survey materials are prepared.

In summary, while the survey profession has established disclosure codes, they are not adhered to by some prominent organizations. There are differing interpretations of what sort of disclosure is adequate. Further, the profession as a whole has not come to grips with the cost and logistics of the disclosure codes it has adopted. As a result, efforts to explore the inner workings of polls remains a difficult, ad hoc business both for those who request information and those who must deal with requests.

Overview of the polls. Based on information from cooperating organizations, it appears that the 1991 polls conducted at the time of the hearings (in the upper part of table 1) shared several features. Random digit dialing (RDD) for household selection seems to have been employed. In most of the polls (eight out of 11) the field work lasted a single evening. Apart from the CBS/*New York Times* use of a panel design, all were independent samples. Most of the field work for the 1991 polls was done over the weekend when Anita Hill and Clarence Thomas and their supporters testified. The Hart-Teeter poll for NBC/*Wall Street Journal* screened for registered voters, whereas all others targeted the broader adult population. The reported response rates were around 50 percent. (The methods of response rate calculation are examined below. The CBS/*New York Times* figures are panel recontact rates.)

The 1992 "year-after" polls used RDD for household selection. Unlike the 1991 polls, the Hill-Thomas questions in 1992 were embedded in omnibus surveys covering a range of topics. None of the 1992 polls had field work completed in a single evening. The one reported response rate for the three polls was around 30 percent.

Respondent selection and callbacks. The polls employed three different procedures for selecting respondents within households. The *Los Angeles Times*, Hart-Teeter, Hart-Breglio, and ICR Survey Research (for Princeton Survey Research Associates) used versions of the "next/last birthday" method (Salmon and Nichols 1983). The CBS/*New York Times* survey used the Troldahl/Carter method (1964). Gallup used a unique procedure that asks for the youngest male or oldest female, and then substitutes if that person is not available or refuses. Gallup also imposes a gender quota so that roughly half of the interviews are taken with men and half with women. Though information about the specific Hill-Thomas poll was not provided, a 1993 Louis Harris and Associates description of its national telephone sample (Harris and Associates 1993) states that the youngest male/oldest female respondent selection procedure is used in some studies and a random selection procedure is used in others.

The methodological literature offers different views on the desirability of these respondent selection methods. The "next/last birthday" method selects respondents by asking the phone answerer which household member had the last birthday, or which one will have the next birthday. Hart-Teeter, Hart-Breglio, the *Los Angeles Times*, and ICR modified this procedure with a probe for the phone answerer ("Well, of the ones you do know, who had the last birthday?") if he or she was unsure about the household member with the last birthday. Groves and Lyberg (1988) provide one perspective on this general approach. They argue that the birthday procedure produces a nonprobability sample because birthdays are not randomly assigned to individuals within households and because phone answerers do not report birthdays accurately (as the probe employed by several organizations suggests). Lavrakas, Bauman, and Merkle (1993) lend empirical support to this view, finding substantial error in reporting of "last birthdays" by household informants and resultant coverage problems. Crespi's influential preelection poll volume (1988), on the other hand, says that the "birthday" method appears to be "theoretically well based" (p. 52).

The Troldahl/Carter method (1964) also evidences within household coverage problems, according to research by Groves and Kahn (1979). Crespi (1988) ranks the method with the Kish (1949) household listing procedure as theoretically advantageous, but notes that a number of pollsters argue that it increases refusals and the need for callbacks in telephone interviews. In this context, the practical advantages of the Gallup substitution selection method for getting interviews quickly are apparent, though the technique has no roots in any statistical theory.

The one-night field periods associated with many of the Hill-Thomas polls made callbacks difficult or impossible. The reported number of callbacks varied among organizations from two (*Los Angeles Times* and Gallup) to "as many as necessary," but in practice only one call was made in many cases. For example, the *Los Angeles Times* reported that two callbacks were made before allocating a phone number to the "no answer" noncontact category, but also maintained an "insufficient callback" category that contained some 2,700 cases. Gallup reported allowing two callbacks, but also noted that the organization typically does not make callbacks during a single-night poll. Louis Harris and Associates' (1993) "national telphone sample" description states that three callbacks are made, and substitute households are selected if refusal conversions fail. Most of the organizations contacted for this review could not or did not provide an average number of actual callbacks for the polls in question.

RESPONSE RATES

Methods of calculating response rates differ among the organizations who cooperated with requests for information for this article. Figuring out how to compare them presents a dilemma. In theory, the response rate should express the number of people interviewed as a proportion of those eligible to be interviewed. As is well-known, determining eligibility is not straightforward in telephone surveys. Many dialed numbers are not answered, and the eligibility of persons attached to the numbers is unknown (if, in fact, they are working numbers). At the respondent level, selected persons may not be available, or may refuse initially to be interviewed, requiring callbacks. The shorter the field period, the less one is able to judge whether unanswered numbers might produce an eligible respondent if called again, or if initial refusals are firm refusals. Data collection completed in one night is certain to leave many unanswered questions about eligibility and also fail to produce interviews with reluctant respondents. Any discussion of eligibility is further complicated by respondent selection methods that permit substitution. In assessing the Hill-Thomas polls, we encounter both the problem of eligibility definition and substitution.

Looking again at table 1, Gallup reports the highest response rate (54 percent) for a four-night poll (October 10–13, 1991) and a lower rate (47 percent) for the single-night poll (October 14, 1991). The response rate reported by Gallup, however, is a number primarily derived for internal use and is not a ratio of completed interviews to sample cases in which an interview could have been completed. The Gallup response rate is derived by multiplying three proportions: the "contact rate" (the proportion of working household numbers at which a contact was made), the "cooperation rate" (the proportion of contacted households that provide a willing respondent, and the "completion rate" (the proportion of cooperating households that provide a completed interview). Recalling that Gallup substitutes when the selected respondent is unavailable or refuses, the "cooperation rate" figure is certain to be higher than a similar figure for organizations that do not substitute.

The *Los Angeles Times* reported a response rate of 52 percent. This figure is computed as a ratio of completed interviews to completed interviews plus refusals, unavailable respondents, and other noninterviews (e.g., incomplete interviews, language problems). The denominator of the response rate does not include, however, an estimate of eligible respondents at telephone numbers that received "insufficient callbacks" or where residency status was undetermined. In the unlikely event that all of these telephone numbers produced eligible respondents, the response rate would be around 23 percent for a two-night poll (October 12–13). Since at least some of the "insufficient callback" and "residency undetermined" cases were apt to produce eligible respondents, however, it is fair to say that the response rate for this poll fell in the range between 23 and 52 percent. On September 3–5, CBS/*New York Times* completed the first wave of its panel on the Hill-Thomas issue. The response rate for this effort ranged between 56 percent and 71 percent, depending on how many of the "ring, no answer" cases were eligible. A separate panel of African-American respondents had a response rate ranging from 62 to 68 percent on the first wave, also conducted September 3–5.

The *Los Angeles Times* and CBS/*New York Times* were the only organizations able to provide call record data permitting calculations such as that just described for the 1991 polls. The reported response rate for the ICR/Princeton Survey Research Associates 1992 poll (using a formula like Gallup's) ranged from 27 to 36 percent, depending on assumptions about the "contact rate" (how many of the noncontacts were considered eligible). Gallup's reported response rate for their 1992 poll was 42 percent.

In summary, what is known about respondent selection and callbacks in the Hill-Thomas polls suggests that these efforts often involved substantial departures from probability sampling due to selection methods employed and constraints on callbacks. The cause of both the selection methods and the lack of callbacks appears to be the breaking news–determined field work constraints within which the polling organizations must work. Because of the extraordinary time pressure, respondent selection methods that will obtain an interview with someone on the first call are to be preferred. Significant trade-offs are made quickly and routinely between coverage and nonresponse errors. It is left to quota controls and weights to adjust the resulting samples. Reported response rates are sometimes not constructed through formulae that one might expect, having consulted the methodological literature (e.g., Groves 1989). Because of the short field work periods, there is more uncertainty than in the usual telephone survey about the status of noncontacts, and reported response rates may not reflect that uncertainty. They are likely to be high estimates. Despite the efforts to reduce it through respondent selection methods and substitution, nonresponse is a substantial issue in these fast reaction polls.

The CBS/*New York Times* panel design offers an interesting alternative to doing a fast reaction poll "from scratch." Recontacting panel members obviates the need to screen for working household numbers. Additionally, since much is known about respondents from the survey's first wave, characteristics of respondents and nonrespondents in subsequent waves can be compared. Of course, the problem of panel attrition is still present—especially in a fast reaction setting—as one can note from the recontact rates for the CBS/*New York Times* poll in table 1.

INTERVIEWING

Going beyond the sampling considerations displayed in table 1, there are questions about the conduct of interviewing that are interesting to examine in the Hill-Thomas case. Insofar as can be judged, none of the organizations that conducted polls dealing with this matter attempted to measure or control the influences of interviewer race or gender on responses. Both race and gender of interviewers have been shown to affect responses to opinion items that are associated with those social categories, although the nature of the relationships is not fully understood (see Groves 1989; Kane and Macaulay 1993; Lewptow, Moser, and Pendleton 1990; Schaffer 1980; Schuman and Converse 1971). In the Hill-Thomas matter, both race and gender were involved in the controversy and the interviewer's perceived race or gender may have had an impact on responses. One might hypothesize, for example, that after Clarence Thomas's angry denouncement of the Senate's "high tech lynching" that black respondents interviewed by white interviewers may have been more likely to show solidarity with Thomas (see Mansbridge and Tate 1992). In this connection, it is worthwhile to note that only a few organizations (CBS/*New York Times*, *Los Angeles Times*, and, reportedly, Gordon Black) attempted oversamples of African-Americans that would have been required to test this hypothesis.

QUESTIONNAIRE DESIGN

The issues of question wording, order, and context similarly provide substantial grounds for speculation in the Hill-Thomas polls. All of the polls for which information could be obtained, for example, asked early in the interview—following the format common in preelection polls—whether or not Clarence Thomas should be confirmed. This item was followed by the questions on Thomas's and Hill's credibility (see App. A). The order of these items is set so as to avoid the possibility of a question order effect on the confirmation item. The chosen order, however, may have affected responses to the credibility items instead, if respondents reported favoring confirmation and wanted to give consistent credibility responses.

Perhaps more important than question order is the framing of the credibility questions themselves. As seen from a perusal of Appendix A, many of the items force choices, using phrases such as "just your best guess" or "if you had to choose" or "based on what you have heard so far." Respondents mainly had to volunteer "don't know" responses. Such question construction seems to be essential in fast reaction polls, since their aim is to get some public assessment of a breaking news event. "Don't know" responses, though they may be more accurate statements of sentiment concerning rapidly changing events, do not appear to have much news value. But the strategy of forcing choices is likely to favor the status quo option—in this case, Clarence Thomas—since respondents are more likely to have heard of it. The fact that Thomas appeared on national television for considerably more time than Hill (and also appeared during "prime time") to testify in the Senate hearings accentuates this possibility.

Thus, the apparent strong support for Thomas in the 1991 polls is open to question, as is the apparent shift in support from Thomas to Hill in the 1992 polls. The 1991 and 1992 poll findings on credibility arrayed for reference in Appendix B may be more indicative of a process of opinion formation rather than opinion change. Fast reaction polls and the way they are reported by their media sponsors often give the impression that the public has opinions when it may not. As Kathleen Frankovic pointed out in her 1993 AAPOR presidential address (Frankovic 1993), pollsters must consider the fact that they sometimes make as well as measure opinions.

Conclusion

This article has highlighted the difficulty in documenting the details of public opinion polls that are the primary source of information about public sentiment in historical accounts of public events. The search for information about the Hill-Thomas polls suggests that norms of professionalism are not institutionalized generally in the polling community. In those organizations that do wish to comply with disclosure requests, costs and logistical difficulties, as well as unsupportive clients or managements, may make fulfilling the requests difficult. This state of affairs appears to be particularly true for information about sampling and field work. A possible solution to the difficulties faced by both those seeking information and those responding to requests is a routinized deposit of such information in an archive such as the Roper Center. Organizations that employ CATI systems may find that archiving call records—with appropriate safeguards for respondent anonymity—is not an onerous procedure.

Based on the available information, this article has raised questions about the trade-offs involved in fast reaction polls. Substantial nonresponse, or at least substantial indeterminacy in constructing response rates, was illustrated. Methods of respondent selection were discussed, and ongoing disagreements about the value of different methods were noted. It seems clear from this discussion that substantially more methodological work needs to be done to assess selection methods that are routinely used by survey organizations.

In addition to the time pressure in fast reaction polls that affects response rates, the pressure on polling organizations to collect opinions about breaking news events leads to questioning strategies that suppress public ignorance or indecision. News organizations appear to need opinions, not "don't knows." In addition, news reports on polls do not treat the ambiguity inherent in an event such as the Hill-Thomas dispute. Though it was possible to believe or disbelieve both

Hill and Thomas, the story line based on forced-choice questions did not give credence to these possibilities. The richness and subtlety in public sentiment that may be measured in public opinion polls often is not mentioned in news accounts, and this is especially true in the case of fast reaction polls.

When polls in 1992 suggested that public opinion had shifted in favor of Anita Hill, the inference of a change was based on a comparison with the "instant"-polls of a year earlier. Scrutiny of the "first draft of history" based on those efforts has pointed to the need to better safeguard the information upon which historical judgments are based and to the dangers inherent in collecting the information under conditions that threaten its validity.

Appendix A

Hill-Thomas Truth/Belief Items

CBS/NEW YORK TIMES

[Preceding item(s) (10/9/91): "What do you think now? Should the Senate vote to confirm Clarence Thomas as a justice of the U.S. Supreme Court, or vote against Thomas, or can't you say?" If the panel respondent gave a different response than that given in September, the respondent was reminded of the different response and asked, "What caused you to change your mind?"]
[Preceding item (10/13/91, 10/14/91): "Is your opinion of Professor Anita Hill favorable, not favorable, undecided, or haven't you heard enough about Anita Hill yet to have an opinion?"]
Question (10/9/91, 10/13/91, 10/14/91): "Anita Hill, a law professor and former aid to Thomas, has charged that Thomas sexually harassed her while he was head of the Equal Employment Opportunity Commission. Thomas has denied the charges. Right now, what do you think—do you think the charges of sexual harassment are probably true, or probably not true?"
[Preceding item (10/13/91, 10/14/91): "In her testimony to the Senate Judiciary Committee do you think Anita Hill told the entire truth about what happened between herself and Clarence Thomas, or was she mostly telling the truth but also hiding something, or do you think she was mostly lying about it?"]
Question (10/13/91, 10/14/91): "Who do you believe *more*—Clarence Thomas or Anita Hill?"

CHILTON FOR ABC/*WASHINGTON POST*

[Preceding question unavailable.]
Question (10/12/91): "Who do you think was the more believable witness—Hill or Thomas?"

[Preceding question unavailable.]
Question (10/13/91): "Just your best guess—do you think the charge against Thomas is true or not true, or don't you know enough to say?"

LOUIS HARRIS AND ASSOCIATES

[Preceding question unavailable.]
Question (10/9–13/91): "Now, if you had to choose, which do you tend to believe most—the charges made by Professor Anita Hill or the denials made by Judge Clarence Thomas?"

GALLUP

[Preceding question unavailable.]
Question (10/10–13/91, 10/14/91): "Anita Hill charges Clarence Thomas with sexually harassing her when she worked for him in the early 1980s. Thomas denies Hill's charges. From what you have heard or read, who do you believe more—Anita Hill or Clarence Thomas?"
Question (10/1–3/92): "I have a question about the Supreme Court confirmation hearings for Clarence Thomas held last year. During the hearings, Anita Hill charged Clarence Thomas with sexually harassing her when she worked for him in the early 1980s. Thomas denied Hill's charges. From what you remember or have heard or read since that time, who do you believe more—Anita Hill or Clarence Thomas?"

LOS ANGELES TIMES

[Preceding item (10/12–13/91): "How serious are Anita Hill's accusations against Clarence Thomas? Are they very serious, somewhat serious, not too serious or not at all serious?" (Accept "Haven't heard enough to say" as a volunteered response.)]
Question (10/12–13/91): "Based on what you've heard so far, do you think Anita Hill's charges of sexual harassment against Clarence Thomas are definitely true, or probably true, or probably false or definitely false?" (Accept "Haven't heard enough to say" as a volunteered response.)
Question (10/12–13/91): "What about Clarence Thomas's denials of sexual harassment charges against him? Based on what you've heard so far, do you think Clarence Thomas's denials are definitely true, or probably true, or probably false or definitely false? (Accept "Haven't heard enough to say" as a volunteered response.)
Question (10/12–13/91): "Which version of events do you tend to believe more: that told by Anita Hill or that told by Judge Clarence Thomas?" (If answer "Hill" or "Thomas," ask: "Are you certain about that?")

GORDON S. BLACK CORPORATION FOR USA TODAY

[Preceding question unavailable.]
Question (10/13/91): "Thomas has been accused of sexual harassment by Anita Hill, and he has denied the charges. Who do you believe is telling the truth?"

HART-TEETER FOR NBC/WALL STREET JOURNAL

[Preceding question (10/15/91): "Do you think Clarence Thomas did or did not receive a fair hearing before the Senate Judiciary Committee?"]
Question (10/15/91): "In general, who do you think was telling the truth about Anita Hill's charges of sexual harassment by Clarence Thomas—Anita Hill or Clarence Thomas?" (Accept only one response. Rotate on every other interview.)

ICR SURVEY RESEARCH FOR PRINCETON SURVEY RESEARCH ASSOCIATES AND U.S. NEWS AND WORLD REPORT

[Preceding item (9/11–15/92): "As you may recall, a year ago Professor Anita Hill accused Clarence Thomas of sexual harassment before the Senate Committee considering his confirmation as a Supreme Court Justice. All things considered, do you think the Senate did the right thing in confirming Clarence Thomas as a Supreme Court Justice or do you think it did the wrong thing in confirming him?"]
Question (9/11–15/92): "Looking back on the charges that Anita Hill made and Clarence Thomas's responses, who do you believe more, Clarence Thomas or Anita Hill?"

HART-BREGLIO FOR NBC/WALL STREET JOURNAL

[Preceding question (9/12–15/92): "The Bush Administration has negotiated a North American Free Trade Agreement with Mexico to eliminate nearly all restrictions on imports, exports, and business investment between the United States and Mexico. Do you favor or oppose this trade agreement with Mexico? If you feel you have not heard enough about this issue yet to have an opinion, please just say so."]
Question (9/12–15/92): "It has now been 1 year since the congressional hearings on the dispute between Anita Hill and Clarence Thomas. In general, who do you think was telling the truth about Anita Hill's charges of sexual harassment by Clarence Thomas—Anita Hill or Clarence Thomas?"

Appendix B

Table B1. Anita Hill–Clarence Thomas: Whom Do You Believe?

	1991						
	Total		Men		Women		
	Pro-Hill	Pro-Thomas	Pro-Hill	Pro-Thomas	Pro-Hill	Pro-Thomas	N
CBS/*New York Times:*							
Which telling truth?							
10/9/91	21[a]	47	18	53	23	42	512
10/13/91	27	54	26	54	28	55	501
10/14/91	27	56	20	63	34	49	497
Whom believe more?							
10/13/91	24	58	22	59	26	57	501
10/14/91	20	60	16	63	34	49	497
Chilton for ABC/*Washington Post:*							
Whom believe more?							
10/12/91	24	46	20	56	28	38	513
Which telling truth?							
10/13/91	37	54	33	59	41	49	506
Gallup:							
Whom believe more?							
10/10–13/91	29	48	27	51	31	46	1,000
10/14/91	27	54	23	55	31	53	778
Los Angeles Times:							
Hill's charges true?							
10/12–13/91	35	47	34	51	41	41	1,264
Thomas's denials true?							
10/12–13/91	33	51	33	53	34	48	1,264
Whom believe more?							
10/12–13/91	37	46	32	52	38	44	1,264

Gordon S. Black Corp. for *USA Today:* Which telling truth? 10/13/91	24	47	22	49	26	45	756
Louis Harris and Associates: Whom believe more? 10/9–13/91	33	56	28	61	37	51	1,254
Hart-Teeter for NBC/*Wall Street Journal:* Which telling truth? 10/15/91	24	40	22	42	27	38	832

	1992						
	Total		Men		Women		
	Pro-Hill	Pro-Thomas	Pro-Hill	Pro-Thomas	Pro-Hill	Pro-Thomas	*N*
ICR Survey Research for Princeton Survey Research Associates and *U.S. News and World Report:* Whom believe more? 9/11–15/92	38	38	27	44	48	33	1,002
Hart-Breglio for NBC/*Wall Street Journal:* Which telling truth? 9/12–15/92	44	34	37	40	50	29	1,506
Gallup: Whom believe more? 10/1–3/92	43	39	39	46	46	36	1,243

NOTE.—Exact question wording appear in App. A.

[a] Percentages do not add to 100 because "don't knows" and other nonresponse is not included here.

References

Anderson, B. A., B. Silver, and P. Abramson. 1988. "The Effects of Race of the Interviewer on Race-related Attitudes of Black Respondents in SRC/CPS National Election Studies." *Public Opinion Quarterly* 52:52–83.

Crespi, I. 1988. *Pre-Election Polling*. New York: Russell Sage.

Frankovic, K. 1993. "Presidential Address: Noise and Clamor: The Unintended Consequences of Success." *Public Opinion Quarterly* 57:441–47.

Gollin, A. 1992. "AAPOR and the Media." In *A Meeting Place: The History of the American Association for Public Opinion Research*, ed. P. B. Sheatsley and W. J. Mitofsky. American Association for Public Opinion Research.

Groves, R. 1989. *Survey Errors and Survey Costs*. New York: Wiley.

Groves, R., and N. Fultz. 1985. "Gender Effects among Telephone Interviewers in a Survey of Economic Attitudes." *Sociological Methods and Research* 14:31–52.

Groves, R., and R. Kahn. 1979. *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.

Groves, R., and L. Lyberg. 1988. "An Overview of Nonresponse Issues in Telephone Surveys." In *Telephone Survey Methodology*, ed. R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg. New York: Wiley.

Harris, Louis, and Associates. 1993. *Louis Harris and Associates' National Telephone Sample*. Available from Louis Harris and Associates, 630 Fifth Avenue, New York, NY 10111.

Hollander, S. 1992. "Survey Standards." In *A Meeting Place: The History of the American Association for Public Opinion Research*, ed. P. B. Sheatsley and W. J. Mitofsky. American Association for Public Opinion Research.

Kane, E., and L. Macaulay. 1993. "Interviewer Gender and Gender Attitudes." *Public Opinion Quarterly* 57:1–28.

Kish, L. 1949. "A Procedure for Objective Respondent Selection within the Household." *Journal of the American Statistical Association* 44:380–87.

Lavrakas, P., S. Bauman, and D. Merkle. 1993. "The 'Last Birthday' Selection Method and Within-Unit Coverage Problems." Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Charles, IL, May 20–23.

Lewptow, L., S. Moser, and B. Pendleton. 1990. "Gender and Response Effects in Telephone Interviews about Gender Characteristics." *Sex Roles* 22:29–42.

Mansbridge, J., and K. Tate. 1992. "Race Trumps Gender." *PS: Political Science and Politics* 25:488–92.

Miller, P., D. Merkle, and P. Wang. 1991. "Journalism with Footnotes: Reporting the 'Technical Details' of Polls." In *Polling and Presidential Election Coverage*, ed. P. Lavrakas and J. Holley. Newbury Park, CA: Sage.

Salmon, C., and J. Nichols. 1983. "The Next Birthday Method of Respondent Selection." *Public Opinion Quarterly* 47:270–76.

Schaffer, N. 1980. "Evaluating Race-of-Interviewer Effects in a National Survey." *Sociological Methods and Research* 8:400–419.

Schuman, H., and J. Converse. 1971. "Effects of Black and White Interviewers on Black Response in 1968." *Public Opinion Quarterly* 35:44–68.

Troldahl, V. C., and R. Carter. 1964. "Random Selection of Respondents within Households in Phone Surveys." *Journal of Marketing Research* 1:71–76.

Turner, C., and E. Martin, eds. 1984. *Surveying Subjective Phenomena*, vol. 1. New York: Russell Sage.

THE POLLS—POLL TRENDS
TRENDS IN MEDIA USAGE

WILLIAM G. MAYER

How do Americans learn about the world around them? What are their primary sources of news? How many of them have regular access to televisions or newspapers, and how often do they use these media? The survey questions presented in this article attempt to answer these questions and, especially, to show how American media usage practices have changed since the late 1930s.

Before presenting an overview of these data, however, a few words of caution are in order, for a substantial amount of controversy surrounds some of the conventional survey questions that are used to measure how people get their news.[1] As we will see, most Americans today say that television is their principal source of news—but a number of scholars have argued that the reality is quite different. While people undoubtedly do spend a large part of their lives watching TV (or doing other things with the TV on in the background), the vast majority of what they watch is entertainment programming rather than news. Even when they do watch the news, television news stories tend to be short, superficial, and disconnected, and elicit a correspondingly low level of intellectual involvement from their audience. As a result, these scholars claim, it is actually newspapers that continue to be the primary vehicle by which most Americans acquire information about the world around them. At a minimum, we can say that the process by which human beings acquire news about the outside world is neither simple nor obvious, and that direct questions posed in conventional mass sample surveys may not be the only (or the best) way to investigate it.

Before one can defend, criticize, or improve upon these survey measures, however, one must first have available the current data—and

---

[1] The literature on this topic is voluminous. Among the most important contributions to it are Graber (1988); Robinson (1977); and Robinson and Levy (1986). For a useful short summary, see Witt (1983).

WILLIAM G. MAYER is assistant professor of political science at Northeastern University, Boston, Massachusetts. The author wishes to thank Lois Timms-Ferrara, John P. Robinson, and Robert Y. Shapiro for their generous assistance in the preparation of this article.

Public Opinion Quarterly Volume 57:593–611 © 1993 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/93/5704-0007$02.50

that is the purpose of this article. I have deliberately cast the net as widely as possible, in order to present a collection of the best question sequences available that deal with how Americans acquire their news and how they relate to the major media of mass communications.

A number of significant trends are apparent in these data, but one overall finding especially deserves to be highlighted. At several points in the past 50 years, American media usage patterns have changed very substantially in a remarkably short period of time. Where several recent works have argued that major social changes generally proceed at a fairly slow and gradual pace (see Davis 1986; Mayer 1992, chap. 8; and Page and Shapiro 1992), media-related behaviors present some striking examples to the contrary. Between 1949 and 1959, for example, the number of Americans who owned television sets jumped from 6 percent to 90 percent. During the 1980s, almost 70 percent of the population acquired videocassette recorders for the first time, while the proportion who subscribed to cable television grew by 30 percentage points.

The data are divided into five sections:

1. *The radio era*. By the time that the pioneers of survey research began doing national sample surveys in the late 1930s, the United States was already well into the radio era. By 1937, according to one early Gallup survey, almost 90 percent of American households owned radios. Yet, if we may judge by the set of questions shown in this section (the question wordings were changed quite frequently, but the general tenor of the results is consistent), radio never established itself in the public mind as the nation's principal source of news. When asked where they got most of their news, Americans tended to mention newspapers and radio in about equal proportions.

The one exception to this pattern—which presents a fascinating parallel with America's more recent experience in the Gulf War—is that World War II apparently increased the country's desire for more immediate, up-to-the-minute news reports, and thus provided a substantial (though perhaps short-lived) boost to the popularity of radio. Between 1939 and 1944, according to the Gallup polls, more and more Americans said that they listened regularly to radio news broadcasts, that they learned about important events first on the radio, and that they preferred getting their news from radio.

2. *The rise of television*. The public debut of television is generally said to have taken place at the 1939 World's Fair in New York City. Its full-scale development was delayed by America's entry into World War II and by the absence of FCC-approved technical standards (see Barnouw 1990, chaps. 2 and 3), but the nascent television industry gradually began to acquire momentum in the mid-1940s. In one of the first polling questions ever asked about television, which appeared in

a December 1945 Gallup survey, 11 percent of the population still did not know what television was, and only 22 percent had ever seen one in operation. As late as May of 1950, the Gallup poll could still meaningfully ask its respondents if they had ever seen a television—and find that fully 32 percent had not.

The best single measure of television's enormous growth in postwar America is the percentage of the population who said they had a television set in their home. In 1949, only 6 percent of the Gallup sample said they had a television. But this jumped to 23 percent in 1951, to 47 percent in 1952, to 76 percent in 1955, to 90 percent in 1959, and then to 97 percent in 1966. In just 10 years, in other words, television went from being a marvel and a curiosity to being a shared and common part of our national life.

If television was becoming a fixture in American homes during the early and mid-1950s, it is not as easy to determine whether or how it was being used as a source of news. So far as I can determine, no one asked a national sample what their principal source of news was during this period.[2] We can perhaps get some purchase on this issue, however, by looking at two surveys that asked parallel questions about how respondents used television, newspapers, and radio. In July 1952 Gallup found that 72 percent of its sample said they had read about the Republican national convention in the newspapers and 58 percent had listened to part of it on the radio, but only 50 percent had seen any of it on TV. (Another series of questions on the Democratic convention shows a similar ordering.) In a 1954 AIPO-NORC study, 71 percent said they read a newspaper almost every day, 51 percent listened to radio news programs almost every day, and 39 percent watched televised news on an almost daily basis.

The final item in this section is perhaps the most frequently cited survey question on American media usage patterns, developed by the Roper Organization for a series of surveys conducted for the Television Information Office.[3] When the question was first posed in 1959, 51 percent of the people said they got "most of [their] news about what's going on in the world today" from television, putting it substantially ahead of radio (34 percent) but still slightly behind newspapers (57 percent). (The percentages sum to more than 100 percent because multiple answers were permitted.) But by 1963, television had inched ahead of newspapers, and it continued to widen its advantage through the 1970s and 1980s. By 1992, 69 percent said television was their principal source of news, 43 percent named newspapers, and radio had shrunk to a mere 16 percent. Since Carter and Greenberg (1965) have criticized this question (wrongly, in my judgment) for permitting respondents to name more than one medium as their major source of information, I have also included a breakdown of these multiple responses.

---

[2] The 1952 and 1956 American National Election Studies did, however, ask respondents where they got most of their information about the election campaigns. For these data, see sec. 5 below.

[3] The Television Information Office closed in 1989. The 1990, 1991, and 1992 Roper surveys were sponsored by the Network Television Association and the National Association of Broadcasters.

3. *Newspapers.* Though the American public's use of newspapers has clearly changed a lot less than its reliance on radios or television, there have been some changes in newspaper reading habits over the past few decades. And as several questions in this section indicate, the general trend clearly has been in a downward direction. The most dramatic evidence of this comes from a NORC question first asked in 1967 and then incorporated into the General Social Surveys. In 1967, 74 percent of the American public said that they read a newspaper every day. By 1991, this had fallen to 51 percent. Meanwhile, the number who said they had read the paper once a week or less rose from 15 percent to 28 percent.

4. *Time use studies.* Some of the most interesting work on how the mass public use media has come from a series of studies designed to investigate the general question of how Americans use time (see esp. Robinson 1977; Robinson, Andreyenkov, and Patushev 1988). The general methodology of these studies is to ask respondents to fill out a "time diary" that shows what they were doing at each moment during the previous day. When aggregated across a national sample, such diaries provide remarkably detailed evidence about the texture of everyday life in the United States.

Major time-use studies were conducted in 1965, 1975, and 1985; but differences in survey methodology and the sampling universe make it difficult to compare time-use patterns across the decades. In this article, I have chosen only to reproduce a table originally compiled by John P. Robinson that shows how much time Americans between 18 and 64 years of age spent on various leisure-time activities. As can be seen, the total amount of leisure time available to the average American increased by almost 6 hours between 1965 and 1985—but almost all of this additional time was absorbed by increased television viewing. The time devoted to reading, by contrast, dropped from 3.7 to 2.8 hours per week.

Robinson's data, it should be noted, measure the time devoted to watching television programs of all kinds, not just the news. In his original 1965 survey, he actually found that on a typical day, more Americans read a newspaper than watched a television news program.

The figures shown here are also limited to the respondent's primary activity and thus exclude those occasions when people have the television (or radio) on while performing other activities. For comparative purposes, I also include two questions from the General Social Survey that asked respondents to estimate how many hours per day they usually spent watching television and listening to the radio. As can be seen, the NORC figures are substantially larger than the time-use study numbers, partly because they include secondary viewing and listening and partly because, according to Robinson, people generally overestimate the amount of time they devote to any particular activity.

5. *Media in campaigns and elections.* The questions in this section all relate specifically to the use of media in the context of election campaigns. In many respects, they reinforce some of the trends discussed earlier, especially the decline of radio, the growth of television, and the slow falloff in newspaper reading. Yet, as a series of Roper questions in this section indicate, the public's use of the mass media clearly varies with the type of office being contested: television's dominance grows greater as the size of the voting constituency increases. In the national campaigns waged for the presidency, about 75 percent of Americans say that they learn most about the candidates from television, as compared to only 20 percent who mention newspapers. In statewide elections, such as those for governor or senator, television's margin is appreciably smaller: 50 percent for TV versus about 30 percent for newspapers. In local elections, a narrow plurality of Americans still say that newspapers are their principal source of information—and almost 10 percent say they learn most from conversations with other people.

The last item in this section is included as a reminder that, even in an era when so much attention is focused on the power of television, Americans still do communicate in face-to-face, unmediated ways. After the 1988 presidential election, 24 percent of the adult population remembered having been contacted by a party campaign worker. Equally significant is that, for all the talk of party decomposition, there is no evidence that the reach of "personal media" has declined. Indeed, the proportion of the population contacted by a party worker was significantly higher in the presidential elections of the 1980s than it had been in the supposedly more partisan 1950s.

All of the Gallup, NORC, and SRC/CPS data reported here, and all of the pre-1950 Roper results, have been verified from the actual codebooks and frequency listings. The Gallup, NORC, and Roper studies are available in the archives of the Roper Center for Public Opinion Research at the University of Connecticut. The SRC/CPS data are made available through ICPSR. Post-1950 Roper figures are based on reports published by the Roper Organization and/or by the Television Information Office.

The number of cases on which these results are based is reported in the tables. The only exception is the Roper/TIO studies, which generally drew on a sample of about two thousand respondents. "Don't know" responses have been included in percentaging all results but are only reported when they equal 1 percent or more of the total.

Sources

AIPO—American Institute for Public Opinion (Gallup)
NORC—National Opinion Research Center, University of Chicago
NORC-GSS—NORC's General Social Surveys
ROPER—Roper Organization
ROPER (for TIO)—Roper Organization surveys for the Television Information Office
SRC/CPS—Survey Research Center/Center for Political Studies, University of Michigan, American National Election Studies

References

Barnouw, Erik. 1990. *Tube of Plenty: The Evolution of American Television*. New York: Oxford University Press.

Carter, Richard F., and Bradley S. Greenberg. 1965. "Newspapers or Television: Which Do You Believe?" *Journalism Quarterly* 42 (Winter): 29–34.

Cutler, Blayne. 1990. "Where Does the Free Time Go?" *American Demographics* 12 (November): 36–39.

Davis, James A. 1986. *What the GSS Tells Us about Social Change, 1972–1985*. GSS Technical Report no. 71. Chicago: NORC.

Graber, Doris A. 1988. *Processing the News: How People Tame the Information Tide*. 2d ed. New York: Longman.

Mayer, William G. 1992. *The Changing American Mind: How and Why American Public Opinion Changed between 1960 and 1988*. Ann Arbor: University of Michigan Press.

Page, Benjamin I., and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.

Robinson, John P. 1977. *How Americans Use Time: A Social-Psychological Analysis of Everyday Behavior*. New York: Praeger.

Robinson, John P., Vladimir G. Andreyenkov, and Vasily D. Patushev. 1988. *The Rhythm of Everyday Life: How Soviet and American Citizens Use Time*. Boulder, CO: Westview.

Robinson, John P., and Mark R. Levy. 1986. *The Main Source: Learning from Television News*. Beverly Hills, CA: Sage.

Witt, Evans. 1983. "Here, There and Everywhere: Where Americans Get Their News." *Public Opinion* 6 (August/September): 45–48.

The Radio Era

AIPO: "Do you have a radio?"

Survey Date	*Yes*	*No*	*N*
April 1937[a]	86	14	2,910
March 1939	87	13	3,227
April 1939	88	12	3,092
March 1940	89	11	3,230
July 1942	91	9	2,934
June 1944	93	7	2,955
December 1956[a]	94	6	1,542
February 1957[a]	92	8	1,504
April 1957[a]	96	4	1,655

[a] "Do you have a radio in your home?"

NORC: "Do you have a radio in working order?"

Survey Date	*Yes*	*No*	*N*
October 1945	88	12	3,243
November 1947	91	9	3,524

AIPO: "Do you listen regularly to news broadcasts [on the radio]?"

Survey Date	*Yes*	*No*	*N*
March 1939	62	38	3,166
April 1939	63	37	3,058
March 1940	66	34	3,151
July 1942	71	29	2,915
June 1944	78	22	1,407

ROPER: "From which source do you get the most of the daily news—the newspapers or radio news broadcasts?"

Survey Date	*Radio*	*Newspaper*	*Both Equal*	*Neither*	*N*
March 1941	39	31	26	3	5,211

ROPER: "From which source do you get most of your news—radio or newspapers?"

Survey Date	*Radio*	*Newspaper*	*Both Equal*	*Neither*	*Don't Know*	*N*
July 1941	36	34	26	3	1	5,218
October 1943	42	35	21	2	1	4,910

NORC: "Where do you get most of your news—from the radio or newspapers?"

Survey Date	*Radio*	*Newspaper*	*Both Equal*	*Neither*	*N*
March 1942	38	25	35	1	2,493
July 1942	39	39	17	4	2,579
March 1945	47	35	15	3	2,498

NORC: "Do you have a radio in working order? [If yes:] From which one source do you get most of your daily news about what is going on—the newspapers or the radio?"

Survey Date	*Don't Own a Radio*	*Newspaper*	*Radio*	*Don't Know*	*N*
October 1945	12	30	54	4	3,243
November 1947[a]	9	44	40	7	3,519

[a] Follow-up question read, "Where do you get most of your daily news about what is going on—from the newspapers or the radio?"

ROPER: "From which source do you get most of your news, radio or newspapers?"

Survey Date	*Radio*	*Newspaper*	*Both Equal*	*Neither*	*N*
May 1946	33	39	25	3	2,527

AIPO: "Would you prefer to get national news from a daily newspaper, or from the radio?"

Survey Date	*Radio*	*Newspaper*	*Both/Neither*	*N*
May 1937	42	52	6	2,840

AIPO: "Do you prefer to get your national and foreign news over the radio or in the daily newspaper?"

Survey Date	*Radio*	*Newspaper*	*Both/Neither*	*N*
March 1939	50	45	5	1,415
April 1939	55	42	3	1,502
March 1940	51	44	4	1,541
July 1942	62	34	4	1,411
June 1944	58	33	10	1,409

AIPO: "When an important event happens, do you usually learn about it first on the radio, or from a newspaper?"

Survey Date	*Radio*	*Newspaper*	*Both/Neither*	*N*
March 1939	68	30	2	1,432
April 1939	73	26	1	1,518
March 1940	71	27	2	1,549
July 1942	81	17	1	2,804
June 1944	82	15	3	2,930

AIPO: "In case the facts in a news broadcast do not agree with those in the newspaper, are you more likely to believe the radio or the newspaper?"

Survey Date	*Radio*	*Newspaper*	*Don't Know*	*N*
March 1939	35	42	22	1,477
April 1939	38	43	19	1,576
March 1940	38	39	22	1,574
July 1942	40	38	23	2,847
June 1944	38	38	23	2,912

The Rise of Television

AIPO: "Do you know what television is? [If yes:] Have you ever seen a television set in operation?"

Survey Date	*Have Seen TV*	*Know What TV Is, but Have Not Seen One*	*Don't Know What TV Is*	*N*
December 1945	22	67	11	3,029

AIPO: "Have you ever seen a program of any kind on a television set?"

Survey Date	*Yes*	*No*	*N*
May 1949	52	48	2,579
May 1950	68	32	2,837

AIPO: [Asked only of those who had seen a television program.] "Where is the set located on which you saw that television broadcast?"

		%	*N*
May 1949	Own home	9	681
	Other home	36	
	Bar or restaurant	29	
	Store	19	
	Other	9	

AIPO: "Do you have a television set in your home?"

Date	*Yes*	*No*	*N*	*Date*	*Yes*	*No*	*N*
May 1949[a]	6	94	1,292	August 1954[d]	59	41	1,618
May 1950[a]	14	86	1,317	August 1954[d]	61	39	1,553
August 1950[a]	15	85	1,328	September 1954	63	37	1,457
January 1951[a]	23	77	1,360	March 1955	72	28	1,504
July 1952[b]	42	58	2,981	April 1955	72	28	1,530
August 1952[b]	45	55	3,026	May 1955	74	26	1,481
August 1952[b]	45	55	3,100	June 1955[a]	72	28	1,461
September 1952[b]	46	54	2,445	June 1955	74	26	1,547
September 1952[b]	44	56	3,039	August 1955	72	28	1,499
September 1952[b]	48	52	2,571	August 1955	74	26	1,495
November 1952[c]	47	53	2,991	September 1955	73	27	1,444
February 1953[b]	49	51	1,593	October 1955	76	24	1,498
April 1953[b]	46	54	1,598	August 1956	80	20	2,166
April 1953[b]	48	52	1,518	November 1956	81	19	1,501
May 1953[c]	45	55	1,555	November 1956	82	18	1,501
June 1953[b]	52	48	1,544	December 1956	83	17	1,540
August 1953	50	50	1,605	February 1957	85	15	1,504
September 1953	50	50	1,513	April 1957	85	15	1,655
November 1953	52	48	1,487	August 1957	84	16	1,539
November 1953	57	43	1,510	September 1957	85	15	1,514
December 1953	56	44	1,471	July 1958	88	12	1,510
February 1954	58	42	1,456	April 1959	90	10	1,688
March 1954	59	41	1,507	June 1964[e]	95	5	1,624
April 1954	62	38	1,474	January 1966	97	3	1,545

[a] "Do you happen to have a television set in your home?"

[b] "Do you happen to have a television set in your own home, or not?"

[c] "Do you own a television set?"

[d] "Do you have a television set in the house?"

[e] "Do you have a television set in your home, or not?"

NORC: "Do you have a television set in your home? [If yes:] About how much time a day do you usually spend watching *news* programs on television?"

		%	*N*
November 1950	Don't own a TV	74	1,266
	Own one, but don't watch news	3	
	Watch 15 minutes or less	8	
	Watch 16–30 minutes	8	
	Watch 31–60 minutes	4	
	Watch more than 1 hour	2	

AIPO: "Did you happen to pay any attention at all to the Republican/Democratic party convention held in Chicago last week? [If yes:] Did you happen to see any part of the Republican/Democratic convention on television? Did you happen to listen to any part of the Republican/Democratic convention on the radio? Did you happen to read anything about the Republican/Democratic party convention in the newspapers?"

July 1952	*Republican Convention*	*Democratic Convention*
Saw on television	50	47
Heard on radio	58	52
Read in newspaper	72	61
N	3,002	1,966

AIPO-NORC: "Do you read a daily newspaper every day, almost every day, or only occasionally?"

Date	*Every Day*	*Almost Every Day*	*Occasionally*	*Almost Never (Volunteered)*	*N*
1954	53	18	21	8	4,924

AIPO-NORC: "Do you have a television set in the house? [If yes:] Do you listen to news programs on television every day, almost every day, or only occasionally?"

Date	*Every Day*	*Almost Every Day*	*Occasionally*	*Almost Never (Volunteered)*	*Don't Own a Television*	*N*
1954	24	15	17	4	40	4,867

AIPO-NORC: "Do you listen to news programs on the radio every day, almost every day, or only occasionally?"

Date	*Every Day*	*Almost Every Day*	*Occasionally*	*Almost Never (Volunteered)*	*Don't Own a Radio*	*N*
1954	34	17	29	16	4	4,917

AIPO: "Do you have a television set in your home? [If yes:] Did you happen to see any news broadcasts on TV yesterday?"

		%	*N*
February 1957	Don't own a TV	15	1,504
	Own one and saw the news	36	
	Own one, but did not see the news	49	

AIPO: "Did you happen to read any newspaper yesterday?"

		%	*N*
February 1957	Yes	70	1,496
	No	30	

AIPO: "Do you have a radio in your home? [If yes:] Did you happen to listen to any news broadcasts on the radio yesterday?"

		%	*N*
February 1957	Don't have a radio	8	1,495
	Have one and listened to the news	39	
	Have one, but did not listen to the news	54	

ROPER (for TIO): "I'd like to ask you where you usually get most of your news about what's going on in the world today—from the newspapers, or radio, or television, or magazines, or talking to people or where?" [Multiple responses permitted.]

Date	*Television*	*Newspapers*	*Radio*	*Magazines*	*People*
December 1959	51	57	34	8	4
November 1961	52	57	34	9	5
November 1963	55	53	29	6	4
November 1964	58	56	26	8	5
January 1967	64	55	28	7	4
November 1968	59	49	25	7	5
January 1971	60	48	23	5	4
November 1972	64	50	21	6	4
November 1974	65	47	21	4	4
November 1976	64	49	19	7	5
December 1978	67	49	20	5	5
November 1980	64	44	18	5	4
December 1982	65	44	18	6	4
December 1984	64	40	14	4	4
December 1986	66	36	14	4	4
November 1988	65	42	14	4	5
December 1990	69	43	15	3	7
February 1991	81	35	14	4	6
November 1992	69	43	16	4	6

Breakdown of Multiple Responses

Date	*TV Only*	*Newspapers Only*	*TV and Newspapers*	*Newspapers and Others but Not TV*	*TV and Others but Not Newspapers*	*Media Other than TV or Newspapers*
December 1959	19	21	26	10	6	17
November 1961	18	19	27	11	7	15
November 1963	23	21	24	8	8	13
November 1964	23	20	28	8	6	12
January 1967	25	18	30	7	8	10
November 1968	29	19	25	6	5	13
January 1971	31	21	22	5	7	13
November 1972	33	19	26	5	5	12
November 1974	36	19	23	4	6	12
November 1976	36	21	23	4	5	11
December 1978	34	19	27	3	6	11
November 1980	39	21	20	3	5	12
December 1982	41	21	20	3	5	10
December 1984	46	22	15	2	3	11
December 1986	50	22	13	1	2	10
November 1988	44	22	18	2	3	10
December 1990	44	19	22	2	3	10
February 1991	54	11	23	1	4	6
November 1992	44	20	21	2	4	9

Newspapers

AIPO/NORC: "Do you usually read a daily newspaper?"

	Survey Date	*Yes*	*No*	*N*
AIPO	April 1937	91	9	2,910
NORC	November 1947	89	10	3,525
NORC	September–October 1948	90	10	2,510

AIPO: "Do you read a daily newspaper regularly?"

Survey Date	*Yes*	*Only Occasionally*	*No*	*N*
March 1939	80	13	7	3,212
April 1939	80	13	7	3,091
March 1940	80	15	5	3,235
January 1941	82	11	8	3,122
July 1942	74	17	9	2,932
June 1943	79	15	7	1,239
June 1944	78	15	6	2,957

AIPO: "Do you read a daily newspaper regularly?"

Survey Date	*Yes*	*Only Occasionally*	*N*
April 1941	83	17	3,087
February 1953	85	15	1,597
December 1956	82	18	1,542

AIPO: "Do you read a daily newspaper regularly?"

Survey Date	*Yes*	*No*	*N*
February 1957	78	22	1,503
April 1957	79	21	1,640
April 1957	84	16	1,656
July 1957	80	20	1,513
October 1957	83	17	1,562
February 1958	83	17	1,477
May 1958	82	18	1,632

NOTE.—Though the question read to the respondents was identical in the three preceding sequences, the precoded response categories printed on the questionnaire varied as indicated.

AIPO/CPS: "Do you read a daily newspaper?"

	Survey Date	*Yes*	*No*	*N*
AIPO	July 1945	84	16	2,612
CPS	Fall 1974	71	29	2,520
CPS	Fall 1976	72	28	2,400
CPS	Fall 1978	71	29	2,304
CPS	Fall 1980	66	34	1,407

NORC: "I'd like to change the subject and ask you how often you read a newspaper: every day, several days a week, about once a week, or less often than that?"

		%	*N*
October 1964	Every day	72	1,974
	Several days a week	12	
	About once a week	9	
	Less often than that	7	

NORC/NORC-GSS: "How often do you read the newspaper—every day, a few times a week, once a week, less than once a week, or never?"

Date	*Every Day*	*A Few Times*	*Once a Week*	*Less than Once a Week*	*Never*	*N*
1967	74	12	6	4	5	3,093
1972	69	15	8	4	4	1,611
1975	66	16	8	5	4	1,488
1977	62	17	10	7	5	1,527
1978	57	20	10	7	5	1,528
1982	54	22	12	7	6	1,503
1983	56	21	11	8	5	1,599
1985	53	21	13	8	6	1,530
1986	54	20	13	8	6	1,468
1987	55	21	13	7	5	1,459
1988	51	24	12	9	5	988
1989	50	25	13	8	5	1,005
1990	53	22	11	10	5	927
1991	51	22	14	9	5	1,022

NORC: "On how many days this last week did you get around to reading a newspaper?"

		%	*N*
November 1950	5, 6, or 7 days	72	1,270
	3 or 4 days	9	
	1 or 2 days	10	
	Not at all	10	

CPS: "How many days in the past week did you read a daily newspaper?"

	1984	*1986*	*1990*
Seven	39	44	43
Six	4	2	2
Five	5	4	4
Four	4	5	5
Three	7	7	7
Two	8	10	9
One	9	10	11
None	23	18	20
N	2,252	2,174	1,997

Time Use Data

AMERICANS' USE OF TIME PROJECT

Hours per week men and women aged 18 to 64 spent on various leisure time activities, as recorded in time diaries:

Activity	*1965*	*1975*	*1985*
Television	10.5	15.2	15.1
Visiting	6.6	5.5	4.9
Reading	3.7	3.1	2.8
Traveling	2.7	2.6	3.1
Talking	2.6	2.2	4.3
Hobbies	2.1	2.3	2.2
Adult education	1.3	1.6	1.9
Cultural events	1.1	.5	.8
Clubs/organizations	1.0	1.2	.7
Religion	.9	1.0	.8
Sports/outdoors	.9	1.5	2.2
Radio/recording	.6	.5	.3
Thinking/relaxing and listening	.5	1.1	1.0
Total	34.5	38.3	40.1

SOURCE.—Cutler 1990, p. 38.

NORC: "On the average, how many hours a day do you watch television?"

		%	*N*
October 1964	None, or don't own TV	7	1,942
	Less than 1 hour	9	
	1 hours–less than 2 hours	21	
	2 hours–less than 3 hours	28	
	3 hours–less than 4 hours	16	
	4 hours–less than 5 hours	10	
	5 hours–less than 6 hours	4	
	6 hours or more	5	

NORC-GSS: "On the average day, about how many hours do you personally watch television?"

	1975	*1977*	*1978*	*1980*	*1982*	*1983*
0	4	4	6	8	4	6
1	17	21	21	18	21	18
2	27	25	27	24	25	25
3	20	20	19	19	19	20
4	15	13	13	13	14	13
5	8	6	6	9	6	6
6–8	8	7	6	7	8	8
9 or more	2	2	2	3	3	2
Average	3.0	2.9	2.8	2.9	3.0	3.0
N	1,483	1,525	1,529	1,454	1,504	1,596

	1985	*1986*	*1988*	*1989*	*1990*	*1991*
0	5	4	3	3	3	4
1	18	18	19	18	22	18
2	26	27	25	30	28	25
3	20	20	19	17	19	21
4	14	13	14	14	12	14
5	8	7	8	7	7	7
6–8	7	8	9	7	7	8
9 or more	2	4	3	3	2	3
Average	3.0	3.1	3.2	3.0	2.9	3.2
N	1,524	1,466	981	999	926	1,014

NORC-GSS: "Do you ever listen to the radio? [If yes:] On the average, about how many hours a day do you usually listen to the radio?"

	1978	*1982*	*1983*
Don't own a radio/never listen	8	10	10
0	1	0	1
1	36	33	33
2	19	19	18
3	8	10	9
4	7	7	7
5	4	5	4
6–8	10	10	12
9 or more	6	7	7
Average	3.0	3.1	3.0
N	1,523	1,501	1,597

Media in Campaigns and Elections

SRC/CPS: "We're interested in this interview in finding out whether people paid much attention to the election campaign this year. Take newspapers, for instance—did you read about the campaign in any newspapers? [If yes:] How much did you read newspaper articles about the election—regularly, often, from time to time, or just once in a great while?"

	1952	*1956*	*1960*	*1964*	*1968*	*1972*
Yes, did read	79	68	80	79	75	57
Regularly	. . .	. . .	44	40	37	25
Often	. . .	. . .	12	14	13	11
From time to time	. . .	. . .	16	18	19	17
Once in a while	. . .	. . .	7	6	7	4
No, did not read	21	31	20	21	25	42
N	1,708	1,762	1,829	1,449	1,351	1,118

	1976	*1978*	*1980*	*1982*	*1984*	*1986*	*1990*
Yes, did read	73	71	71	72	77	64	65
Regularly	25	. . .	. . .	. . .	. . .	. . .	. . .
Often	14	. . .	. . .	. . .	. . .	. . .	. . .
From time to time	23	. . .	. . .	. . .	. . .	. . .	. . .
Once in a while	11	. . .	. . .	. . .	. . .	. . .	. . .
No, did not read	27	29	29	28	23	36	35
N	2,401	2,265	1,405	1,418	1,943	2,175	1,997

SRC/CPS: "Did you listen to any speeches or discussions about the campaign on the radio? [If yes:] How many programs about the campaign did you listen to on the radio—a good many, several, or just one or two?"

	1952	*1956*	*1960*	*1964*	*1968*	*1972*	*1976*
Yes, listened	70	45	42	48	41	43	45
A good many	. . .	. . .	15	12	12	9	11
Several	. . .	. . .	17	23	16	20	20
One or two	. . .	. . .	10	12	12	14	13
No, did not listen	30	55	58	52	59	57	55
N	1,709	1,760	1,826	1,449	1,342	1,119	2,401

	1978	*1980*	*1982*	*1984*	*1986*
Yes, listened	46	47	46	45	41
A good many	8	10	. . .	10	8
Several	21	21	. . .	20	18
One or two	17	16	. . .	16	14
No, did not listen	53	53	53	55	59
N	2,303	1,406	1,416	1,944	2,176

SRC/CPS: "Did you read about the campaign in any magazines? [If yes:] How many magazine articles about the campaign would you say you read—a good many, several, or just one or two?"

	1952	*1956*	*1960*	*1964*	*1968*	*1972*	*1976*
Yes, read	40	31	41	39	35	33	48
A good many	. . .	. . .	12	10	9	6	12
Several	. . .	. . .	15	16	12	12	20
One or two	. . .	. . .	13	13	15	14	16
No, did not read	60	69	59	61	64	67	52
N	1,706	1,758	1,824	1,449	1,351	1,117	2,400

	1978	*1980*	*1982*	*1984*
Yes, read	25	35	31	35
A good many	4	7	. . .	7
Several	11	16	. . .	16
One or two	10	12	. . .	12
No, did not read	74	65	69	65
N	2,304	1,406	1,418	1,944

SRC/CPS: "Did you watch any programs about the campaign on television? [If yes:] How many television programs about the campaign would you say you watched—a good many, several, or just one or two?"

	1952	*1956*	*1960*	*1964*	*1968*	*1972*	*1974*
Yes, watched	51	74	87	89	89	88	55
A good many	. . .	. . .	47	41	42	30	12
Several	. . .	. . .	29	34	34	40	26
One or two	. . .	. . .	11	13	13	18	16
No, did not watch	49	26	13	11	11	12	45
N	1,656	1,760	1,829	1,449	1,345	1,118	2,496

	1976	*1978*	*1980*	*1982*	*1984*	*1986*	*1990*
Yes, watched	89	69	86	76	86	75	63
A good many	34	12	23	. . .	25	. . .	12
Several	35	29	38	. . .	37	. . .	27
One or two	19	27	24	. . .	24	. . .	24
No, did not watch	11	31	14	24	14	25	37
N	2,402	2,300	1,407	1,418	1,944	2,175	1,997

NOTE.—In 1974, the question read, "A lot of people didn't pay much attention to the election campaigns this year. How about you—did you watch any programs about the campaign on television?"

SRC/CPS: "Of all these ways of following the campaign, which one would you say you got the most information from—newspapers, radio, television, or magazines?"

	1952	*1956*	*1960*	*1964*	*1968*
Newspapers	23	24	23	24	21
Radio	28	11	5	4	4
Television	32	49	60	58	62
Magazines	5	5	4	7	5
More than one	3	3	3	4	2
None/don't use any of these media	6	8	5	3	5
N	1,672	1,742	1,813	1,448	1,332

NOTE.—This question was asked after the four preceding questions, and only of respondents who said they had followed the campaign in two or more media. Respondents who said they followed the campaign in only one medium were assigned to that category. Those who answered "No" to all four preceding questions were put in the "None/don't use any" category.

ROPER (for TIO): "During the last election campaign, from what source did you become best acquainted with the candidates running in local elections—like mayor, members of the state legislature, etc.—from the newspapers, or radio, or television, or magazines, or talking to people or where?"

Date	*Newspapers*	*Television*	*Radio*	*People*	*Magazines*	*Other*
January 1971	41	27	6	19	1	5
November 1972	41	31	7	23	1	5
November 1974	41	30	8	14	1	5
November 1976	44	34	7	12	2	6
December 1978	45	39	10	15	1	7
November 1980	36	44	6	11	2	5
December 1982	39	37	6	15	1	7
December 1984	41	39	6	12	1	5
December 1986	33	43	4	13	1	5
November 1988	40	36	5	11	1	5

ROPER (for TIO): "From what source did you become best acquainted with the candidates for the United States House of Representatives from this district—from the newspapers, or radio, or television, or magazines, or talking to people or where?"

Date	*Newspapers*	*Television*	*Radio*	*People*	*Magazines*	*Other*
November 1974	35	40	6	8	1	5
December 1978	38	48	8	10	2	6
December 1982	31	46	5	9	2	7
December 1984	33	49	4	7	1	4
November 1988	35	43	3	5	1	5

ROPER (for TIO): "From what source did you become best acquainted with the candidates running in statewide elections—like U.S. Senator and Governor—from the newspapers, or radio, or television, or magazines, or talking to people or where?"

Date	*Newspapers*	*Television*	*Radio*	*People*	*Magazines*	*Other*
January 1971	29	51	6	10	2	4
November 1972	39	49	7	9	1	3
November 1974	33	48	6	6	1	3
November 1976	35	53	5	6	1	3
December 1978	39	55	8	8	2	5
December 1982	29	53	6	7	1	5
December 1984	33	54	6	6	2	3
November 1988	31	50	5	6	1	4

ROPER (for TIO): "And what about candidates in the national election—for President and Vice President?"

Date	*Newspapers*	*Television*	*Radio*	*People*	*Magazines*	*Other*
November 1972	26	66	6	5	5	2
November 1976	20	75	4	3	5	1
December 1984	19	75	4	3	3	2
November 1988	20	70	4	3	3	2
November 1992	18	74	6	7	2	1

SRC/CPS: "As you know, the political parties try to talk to as many people as they–can to get them to vote for their candidates. Did anyone from one of the parties call you up or come around and talk to you about the campaign this year?"

	Presidential Election Years				*Congressional Election Years*		
Date	*Yes*	*No*	*N*	*Date*	*Yes*	*No*	*N*
1952	12	88	1,709				
1956	17	82	1,752				
1960	21	77	1,823				
1964	26	74	1,448	1966	25	75	1,285
1968	26	74	1,345				
1972	29	71	1,119	1974	24	76	2,523
1976	28	71	2,399	1978	30	69	2,292
1980	24	75	1,408	1982	31	68	1,406
1984	24	76	1,941	1986	25	74	2,174
1988	24	76	1,775	1990	19	81	1,997

BOOK REVIEWS

Benjamin I. Page and **Robert Y. Shapiro.** *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences.* Chicago: University of Chicago Press, 1992. 488 pp. $59.95 (cloth); $19.95 (paper).

STANLEY FELDMAN
State University of New York, Stony Book

The development of survey research made possible the systematic study of public opinion. Since a single survey provides data on the opinions of a large number of people at one point in time, it is not surprising that the dominant mode of analysis in the public opinion field has reflected those characteristics of the data. Although the specific statistical methods may vary, the overwhelming bulk of research on public opinion either describes the distribution of responses to specific questions or estimates relationships among responses to two or more questions. As archives of survey data began to accumulate, with some questions repeated verbatim, the possibility of time series analysis on aggregated responses became possible. The potential of this type of design was probably first demonstrated by research on presidential popularity. In recent years, more investigators have begun to take advantage of repeated surveys to examine the aggregate dynamics of public opinion.

The Rational Public, by Benjamin Page and Robert Shapiro, is the most ambitious study of aggregate public opinion yet produced and one of the most significant contributions to the overall study of public opinion in recent years. The empirical analysis in this book rests on an enormous data base of marginal frequencies from national surveys—as the authors put it, "from *all* published or otherwise available surveys of the American public's policy preferences that we could find, beginning with the first Gallup and Roper polls of 1935 and continuing into 1990" (pp. 42–43). This data base allows them to construct over a thousand time series of varying lengths using identically worded survey questions. The result is an impressive description of trends in public opinion over the last 50 years. There is simply no other research that compares with the detailed description of public opinion that is found here.

This book is not simply an attempt to map out trends in public opinion. Page and Shapiro have much larger goals than this. They want to show that, in the aggregate, public opinion is stable, sensible, and responsive to the social and political environments. How is that possi-

Permission to reprint a book review printed in this section may be obtained only from the author.

ble, given the large volume of research from individual level analyses that appear to show that most people are ill-informed about politics, nonideological, and, very possibly, without opinions on many issues? The authors' answer to this is provided in the first chapter. Whatever may be the vagaries of opinions at the individual level, it is the process of aggregation that makes public opinion intelligible. Although Page and Shapiro reject the extreme claim that many people do not even have opinions on many issues, aggregation does eliminate the noise that is so obvious in individual-level data, producing a picture of largely stable aggregate opinion that is apparently consistent with changes in the social and political environment.

Chapters 3–6 are in many ways the heart of this book. Whatever reaction there may be to the theoretical perspective that Page and Shapiro develop, the description of aggregate public opinion in these four chapters should be required reading for anyone interested in public opinion or American politics more generally. Graphs clearly illustrate trends in repeated survey items and the text adds responses to questions asked less often. The authors frequently point out nuances in responses to related questions that illustrate nicely that the public as a whole differentiates in sensible ways between various types of policy options.

Not only do Page and Shapiro provide a detailed description of public opinion in the United States since World War II (and in some cases before), they also set public opinion in the changing context of American politics and society of the last half century. This is a major strength of these four chapters. Too often, discussions of public opinion present data as if those opinions had materialized out of nowhere. Page and Shapiro are especially careful to relate changes in public opinion to events that help to make sense of those movements. This is best illustrated in the two chapters on foreign policy opinions (coauthored with John Gillroy). So nicely is public opinion located within its political and social context that these chapters also provide an excellent history lesson on international events and American foreign policy from World War II until the present. As much as any analysis in this book, the juxtaposition of public opinion and historical events shows how sensible changes in aggregate public opinion appear to be.

The remaining chapters of the book take up major issues in the study of aggregate public opinion. Page and Shapiro show that, for the most part, trends in public opinion do not differ substantially across social and demographic groups. In part this helps to justify their decision to treat aggregate opinion as a unitary concept. Why this is so is an interesting question in and of itself. The authors offer some possible explanations, but the data cannot provide any direct evidence about this.

Although public opinion appears quite stable in the aggregate, it does show systematic movement. In chapter 8, Page and Shapiro present a rather brief discussion of what is obviously a very complicated problem. Some of this discussion necessarily must be speculative. The effects of social and economic trends on aggregate public opinion may be large over the long term but good time-series data and very careful modeling of the processes are required to draw any firm conclusions. The authors provide an interesting discussion of significant social trends but no evidence of their relationship to public opinion. Their analysis of the effects of the media is on more solid ground, although further analysis is certainly needed.

Finally, Page and Shapiro consider the question of the possible manipulation of public opinion. Given their concern with the overall rationality of public opinion, the issue of manipulation is potentially very important. If opinion is largely dictated by information controlled by those in power, the independent role of public opinion in policy formation would be questionable. Using short case studies as examples, the authors show that there may very well be situations in which the public is manipulated by either the control of information or the presentation of biased information. Ultimately, the question is how much public opinion is controlled in this way. Although Page and Shapiro conclude that this is not a great threat to American democracy, this is an immensely difficult problem and a series of case studies can at best be suggestive.

There are two ways to evaluate this book. As a description of aggregate public opinion and the contexts in which it is observed, Page and Shapiro have produced a resource that is unequaled and is unlikely to be improved on in the near future. This book needs to be read by anyone interested in public opinion or American politics. It should certainly increase the volume of research on aggregate public opinion. Their discussions of opinion change and manipulation raise many important questions for future research. But Page and Shapiro attempt to do more than this. They want to show that, in the aggregate at least, public opinion has the characteristics necessary to be an independent force in a democratic polity. Whether or not you come to believe that aggregate opinion is truly "rational," the authors build a strong case for aggregate public opinion as reasonable and sensible. After much research asserting that public opinion is insubstantial and incoherent, Page and Shapiro's effort is no small accomplishment.

Vincent Price. *Public Opinion.* Communications Concepts, vol. 4. Newbury Park, CA: Sage, 1992. 92 pp. $19.95 cloth; $8.95 paper.

JOHN P. ROBINSON
University of Maryland, College Park

Buy this book! For not much more than a carwash, Stallone movie, or luncheon steak, Price-conscious scholars can have access to this indispensable and insightful guide to the historical and intellectual roots of our profession. Its cost and length also make it highly attrac-

NOTE

ONLY THE FOLLOWING PAGES OF THIS

PUBLICATION HAVE BEEN FILMED

Cov. 1, Pgs: i, ii, 1-48,

77-146

A3610-1A

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

EDITORIAL	
RICHARD R. LAU	An Analysis of "Trial Heat" Polls
PAUL R. ABRAMSON and CHARLES W. OSTROM	Question Wording and Partisanship
DONALD PHILIP GREEN, DANIEL KAHNEMAN, and HOWARD KUNREUTHER	Willingness to Pay for Public Goods
DENNIS O. KALDENBERG, HAROLD F. KOENIG, and BORIS W. BECKER	Response Rates among the Elderly
AÍDA HURTADO	Interviewer Evaluations of Bilingual Respondents
MICHAEL FENDRICH and CONNIE M. VAUGHN	Substance Abuse Underreporting
THE POLLS	Trends: The Rise of the New Media *by William G. Mayer*
BOOK REVIEWS	Reviews *by Vincent Price, Nadia Auriat, Bernard Roshco, Irving Crespi, and Helmut Norpoth*

Spring 1994

BEST AVAILABLE COPY

PUBLIC OPINION QUARTERLY

VOLUME 58 / SPRING 1994 / NUMBER 1

Editorial 1

Articles

An Analysis of the Accuracy of "Trial Heat" Polls during the 1992 Presidential Election 2
RICHARD R. LAU

Question Wording and Partisanship: Change and Continuity in Party Loyalties during the 1992 Election Campaign 21
PAUL R. ABRAMSON and CHARLES W. OSTROM

How the Scope and Method of Public Funding Affect Willingness to Pay for Public Goods 49
DONALD PHILIP GREEN, DANIEL KAHNEMAN, and HOWARD KUNREUTHER

Mail Survey Response Rate Patterns in a Population of the Elderly: Does Response Deteriorate with Age? 68
DENNIS O. KALDENBERG, HAROLD F. KOENIG, and BORIS W. BECKER

Does Similarity Breed Respect? Interviewer Evaluations of Mexican-Descent Respondents in a Bilingual Survey 77
AÍDA HURTADO

Diminished Lifetime Substance Use over Time: An Inquiry into Differential Underreporting 96
MICHAEL FENDRICH and CONNIE M. VAUGHN

The Polls

Trends: The Rise of the New Media 124
WILLIAM G. MAYER

Book Reviews

William A. Gamson, *Talking Politics* 147
VINCENT PRICE

Edouard Cloutier, Jean H. Gray, and Daniel Latouche, *Le virage: L'evolution de l'opinion publique au Quebec depuis 1960* 150
NADIA AURIAT

J. David Kennamer, ed., *Public Opinion, the Press, and Public Policy* 153
BERNARD ROSHCO

Susan Herbst, *Numbered Voices: How Opinion Polling Has Shaped American Politics* 154
IRVING CRESPI

Elisabeth Noelle-Neumann and Renate Köcher, eds., *Allensbacher Jahrbuch der Demoskopie, 1984–1992*, vol. 9 157
HELMUT NORPOTH

Manuscript Referees, 1992–1993 160

Notice to Contributors 161

The articles and reviews in this issue were accepted by the previous editor, Howard Schuman.

Advertising space in *Public Opinion Quarterly* is available, as is rental of its subscriber list. For information and rates, please contact the advertising sales staff, University of Chicago Press, Journals Division, 5720 S. Woodlawn Ave., Chicago, IL 60637. Advertising and list rental are limited to material of scholarly interest to our subscribers.

EDITORIAL

Reading *Public Opinion Quarterly*'s early issues, one is struck by the diversity of contributors: "What Labor Is Thinking" analyzed by John L. Lewis, president of the Congress of Industrial Organizations; "British Public Opinion and Foreign Policy" by Harold Nicolson, member of the House of Commons; "Public Opinion Mechanisms among Primitive Peoples" examined by the anthropologist Margaret Mead; and a French army officer's observations on "National Morale in Wartime France." As these examples suggest, analyses of public opinion in *POQ* did not then depend on survey research. Indeed, at the outset, survey data appeared only occasionally in the *Quarterly*'s articles.

Over time, of course, this changed dramatically. Not only did the proportion of articles drawing on survey data increase, but many of these articles focused on the survey method itself, as opposed to the substance of public opinion. By the late 1940s, about half the articles reporting survey data examined public opinion and half addressed methodological questions—a distribution that is not very different today.

This dual focus on public opinion and survey research arose from the historical coincidence that the leaders in shaping public opinion research were also major contributors to the development of survey methodology. George Gallup, Paul Lazarsfeld, Elmo Roper, Samuel Stouffer, and others played central roles in the growth of both fields. The dividends that accrue from blending a concern for public opinion with an interest in the survey method are illustrated in the work of these pioneers as well as in more recent work such as that of my predecessors as *POQ* editor.

A commitment to both public opinion and survey research does not mean, however, that public opinion research must use the survey method. To the contrary, the field of public opinion is enriched by nonsurvey approaches, including historical studies, laboratory experiments, and qualitative investigations. Likewise, the study of survey methodology can be advanced by a variety of theories and methods (the recent cognitive movement in survey research being a notable example). Thus although the *Quarterly* may no longer be home to the diversity of authors that marked its founding, it continues to be open to the multiplicity of methods that can advance our understanding.

STANLEY PRESSER

AN ANALYSIS OF THE ACCURACY OF "TRIAL HEAT" POLLS DURING THE 1992 PRESIDENTIAL ELECTION

RICHARD R. LAU

Abstract There is little systematic research on the multitude of factors that influence the accuracy of poll results. This article examines six methodological factors directly under the survey researcher's control and two exogenous factors concerning the nature of public opinion as sources of survey error. Data for this study come from 56 "trial heat" polls conducted during the last month of the 1992 presidential election. The most important variables influencing survey accuracy were the number of days a poll is in the field, which increased total accuracy one-half of a percentage point per day; conducting interviews only on weekdays (and thus only during evening hours), which reduced overall accuracy rates by more than 1 percentage point; and conducting a "tracking" poll, which increased accuracy by about 1.5 points. Sample size was not related to accuracy rates. Results also indicated that sampling frames of "likely voters" (relative to "registered voters") tended to overestimate support for George Bush and underestimate support for Ross Perot, that interviewing only on weekdays led to overestimates of support for Bush, and that strict methods of defining a respondent as "supporting" a candidate hurt the two newcomers, Perot and Bill Clinton, more than Bush. In light of these data it is recommended that the common practice of reporting "margins of error" based solely on sample sizes be abandoned as misleading and replaced by a more empirically justifiable measure based more on response rates.

One of the favorite postelection retrospectives played after every presidential campaign is "How well did the polls do in predicting levels of support for the different candidates?" (see, e.g., Broder 1988; Costikyan and Kaplan 1984; Morin 1989; Richburg 1984; Roper 1983). That question seems particularly pointed after the 1993 campaign, with the added volatility caused by the Perot candidacy, and controversies sur-

---

RICHARD R. LAU is professor of political science at Rutgers—The State University. He would like to thank Janice Ballou, Gerry Pomper, Cliff Zukin, and two anonymous reviewers for helpful comments on earlier drafts of this article.

Public Opinion Quarterly Volume 58:2–20 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5801-0003$02.50

rounding wildly fluctuating results due, by some accounts, to poorly explained midcampaign changes in methodology (Brennan 1992; Houston 1992; Morin 1992; Rosenstiel 1992; Taylor and Krane 1993; but see also Hugick, Molyneux, and Norman 1993).[1] President Bush's late surge in the polls, the noticeably greater vigor of his campaigning during the last few weeks of the 1992 campaign, and the general underestimation of the magnitude of the Perot vote may have all been the result of methodological artifacts due to shifts in sampling frames ("registered voters" to "likely voters") or how "undecided" voters are allocated between the candidates. Add to this the overestimation of the magnitude of Clinton's victory by the early exit polls on election day (Public Perspective 1993), and we have a fairly mixed picture of polling results for the 1992 campaign.

Some fluctuation in poll results is inevitable, of course, due to the well-known principles of sampling from any random distribution: the larger the sample, the smaller the sampling error that can be expected. What is much less well known is how other aspects of the survey research process, including choice of sampling frames, response rates, the instability of public opinion, and so on, affect fluctuating poll results. In fact, Buchanan (1986) estimates that observed fluctuations in poll results vary more than twice as much as would be expected by sampling error alone. Converse and Traugott (1986) do an excellent job of listing the variety of factors that could affect accuracy—including sample size, "house" effects, sample composition, interviewer training, questionnaire construction, and so on—but they could report little systematic research that simultaneously considered a host of such factors.

Only Crespi (1988, chap. 9) reports a multivariate analysis that considers the impact of a number of endogenous (sample size, importance of accuracy as a goal for the survey researcher, number of days until the election) and exogenous (primary or general election, whether an incumbent is running for reelection, turnout rates for the election, margin of victory of the winning candidate, percent undecided in the general public) factors, where "endogenous" refers to factors under the survey researcher's control, and "exogenous" refers to characteristics of the environment beyond the researcher's control. The most striking aspect of Crespi's multivariate results is that sample size—the one factor on which we base "margin of error" calculations in reported poll results—did not significantly predict error rates independent of the other variables in the equation.

This article is an attempt to contribute to this line of research. I examine the results of all national "trial heat" polls conducted by eleven different polling organizations[2] between October 1 and election day, as published by the Roper Center in the November/December 1992 issue of *Public Perspective* (see Public Perspective 1992a, 1992b). Using a combination of two different methods to calculate "true" public opinion during this time period (to be described below), I construct measures of "error" or "misestimation" for each poll. The purpose of this article is to develop and test empirically a more comprehensive theory of misestimation or error in "trial heat" surveys.

---

[1] Whether the misunderstanding was the fault of the survey researchers or the media who report their results is another question.

Potential Sources of Error in Trial Heat Polls

Of the many methodological decisions that survey researchers must make in the course of conducting a survey, which ones affect most strongly the accuracy of the results? We could begin by listing clearly worded questions, well-trained interviewers, and the variety of other related factors listed in any textbook on survey methodology. The importance of such factors cannot be downplayed, and they undoubtedly played a role in explaining variance in these data. Unfortunately, I do not have the relevant information to allow these factors to be included in the analysis.

Fortunately, I do have reason to believe they are not as important in explaining variance in my data as they might be in other circumstances. First, the topic of interest—candidate preference in an election—is such a common one that all survey houses have standard, well-rehearsed ways of asking the question.[3] Second, the major polling organizations whose results supply the basic data for analysis should all have standard, high-quality procedures for training and monitoring interviews, data entry, and so on. These survey houses do differ in their standard practices, of course, but such differences are less than would be the case if a number of less experienced survey organizations were included in the data. In any case, I am forced to ignore these factors in the analyses that follow. A more comprehensive list for insuring survey quality cannot ignore these important "textbook" guidelines, however.

---

[2] ABC News (10 polls), CBS News/*New York Times* (seven polls), Gallup (who did 13 polls for CNN/*USA Today*, and five polls for *Newsweek*), Greenberg-Lake/Tarrance Group (five polls), Lou Harris and Associates (three polls), *Los Angeles Times* (two polls), NBC News/*Wall Street Journal* (five polls), Times Mirror Center (one poll), *U.S. News & World Report* (one poll), *Washington Post* (three polls), and Yankelovich Clancy Shulman (one poll).

[3] Not that all survey houses ask this question in an identical manner, however; see Crespi (1988, chap. 5) and the extensive discussion in n. 12. See also Crespi and Morris (1984) on the potential importance of question order when public preferences in more than one race (e.g., a senatorial and a gubernatorial election) are being obtained.

The polls do differ rather sharply on eight factors of which I do have measures, however, and which can become part of the empirical analysis. Six of these factors refer to aspects endogenous to the survey enterprise itself, and two refer to the nature of the public opinion that is being measured.

1. Beginning with choices the survey researcher makes, the first factor is *sample size*. The mathematics of sampling from a random distribution dictate that larger samples should have smaller errors. This is probably the best-known factor affecting the accuracy of survey results. As already noted, however, Crespi (1988) found that sample size was not related to error rates in his multivariate analyses of preelection polling. I will attempt to replicate this surprising finding. The published poll results in my data varied in sample size from 575 to 2,086, which is certainly sufficient to provide another strong test of this hypothesis.
2. The surveys differ in whether their *sampling frame* is "registered voters" or "likely voters." The latter are a subset of the former and, assuming a good algorithm for determining likely voters, should produce more accurate results.[4]
3. The standard formula for determining sampling error in a survey assumes that all sampled respondents are contacted. This of course is far from the truth, and *(non)response rates* should also affect the accuracy of survey results. Unfortunately, survey organizations do not typically publish response rates; indeed, there is not even a standard way of determining them. One fact I do have is the number of days each poll was "in the field." The more days a poll is in the field, the more callbacks are possible for each respondent, and thus the higher the resulting response rate. I hypothesize that the longer a poll is in the field, (the higher the resulting response rate and) the more

---

[4] Defining a "registered voter" is straightforward; but there are many different ways to define "likely voters," and the survey organizations that provided the data for this study undoubtedly differed in the algorithms they apply. These algorithms are often considered "proprietary secrets," however, and again this information is not available for analysis. By necessity, I am forced to treat all methods of selecting "likely voters" as if they are identical. Several survey organizations switch from registered-voter sampling frames to likely-voter frames as election day nears. Hence this dummy variable is somewhat confounded with time. As noted below, however, "Days to election" is included as another predictor in the equation, so that the multivariate results shown in tables 1 and 2 report the effects of sampling frame independent of time.

accurate its results. The number of field days in the current data varies from one to five.[5]

4. A potentially insidious trend that appeared in 1992 was the popular availability of *tracking polls*. Long the staple of campaign consultants, tracking polls were conducted throughout this period by the Greenberg-Lake/Tarrance organization, by Gallup for much of this period, and by ABC news during the final week of the campaign. In tracking polls, a relatively small random sample is interviewed daily with the aim of continuously monitoring the effects of new campaign tactics or exogenous events that occur during the course of the campaign. Taken on a daily basis, they combine the worst of small sample sizes and inevitably low response rates. When the results of 3 or 4 days of such tracking are averaged together, however—as they typically are when reported in the press—they may in fact provide more accurate results than a standard poll. This hypothesis is based on the common practice of "rolling" nonrespondents from one day's polling back into the sample for the next day. This rolling procedure does not go on ad infinitum, but it may go on long enough to ultimately produce more representative samples than is possible with a 2- or 3- or even 4-day poll. The tentative nature of this hypothesis should be noted, but the data are available, and I will test it.
5. Many polling organizations, in order to maximize the proportion of people at home when a call is made, interview only during the early evening hours on *weekdays*, but all day on weekends. When the timing of a survey is such that interviewing is only conducted on weekdays, this practice results in people who work evenings being effectively excluded from the sample. My fifth hypothesis is that surveys conducted only on weekdays will have higher error rates than those conducted on both weekends and weekdays.[6]
6. Survey respondents differ not only on which candidate they support (or whether they claim to support any candidate at all), but also on how strongly they support their favorite candidate (if they have one). Survey houses in turn differ in how strongly

---

[5] There are other ways to increase response rates besides increasing the number of callbacks (see Bradburn 1992), including interviewer training in which the importance of high response rates is stressed; attempting to recontact "refusals" and convert them to completed interviews; assigning the best, most experienced interviewers to any such "conversion" attempts; and scheduling appointments for later interviews.

[6] Interviewing only on weekends would also produce sampling errors. Because there are many fewer weekend days than weekdays, however, this rarely happens. It did not occur at all in the polls I examined.

they probe for different levels of support, and whether they count a respondent who "weakly supports" or "leans toward" a candidate as an actual supporter in their published results. My sixth hypothesis is that survey houses that more actively probe for different levels of support for candidates (and are consequently strict or conservative in reporting levels of support for each candidate) will have lower overall error rates.

7. Turning to public opinion itself, it is only reasonable to predict that the less crystallized opinion is, the more difficulty any survey should have in estimating it. A good estimate of lack of crystallization in public opinion is the *proportion of the sample that cannot choose* between the viable candidates. The larger the percentage of "undecideds" in actual public opinion, the less accurate any poll should be.
8. A related measure of opinion crystallization is time, or rather *closeness of the election* itself. Presumably, the closer the election, the more certain the public should be about which candidate they are going to support. Crespi (1988) found that the number of days until the election was the strongest predictor of error rates, although, as will be revealed below, his method of determining "errors" differed in important ways from mine.

Method

DATA

Public Perspective (1992a, 1992b) reports the support for Bush, Clinton, and (after October 1) Perot in 140 national surveys conducted between August 31 and November 2.[7] Many of these 140 surveys do not involve completely separate samples, however, because many of them were tracking polls. For example, the Greenberg-Lake/Tarrance Group conducted approximately 250 interviews every Sunday, Monday, Tuesday, Wednesday, and Thursday during this period. After each day's polling, new results would be released, based on the last four day's interviews. Thus the 250 interviews conducted each day would be included in four different poll reports. To remove this redundancy, I eliminated from my sample all polls involving overlapping respondents. Thus, for example, the Greenberg poll from October 27–30 was part of my final sample, but the preceding three polls and the following three polls were not included. This results in 76 independent polls that report support for Bush and Clinton, and 56 polls—a subset of the 76—that report support for Perot as well (those conducted after he rejoined the race). Because the results are essentially identical with either sample, I report here data from only the 56 polls that asked for a choice between Bush, Clinton, and Perot.

---

[7] These numbers do not sum to 100 percent because of "undecided" voters. Thus the raw data that I employ are the percent support for each candidate in the poll's total sample—not the percent support for each candidate among those expressing a preference. In calling representatives of the different survey houses to tie up a few loose ends for this article, I learned from Humphrey Taylor of Louis Harris and Associates that *Public Perspective* incorrectly reported the results of the final Harris poll. The errors were slight, however—support for Bush, Clinton, and Perot were all 1 percent too high—and because they could not possibly have affected the results very much, I have not corrected them in my data set.

OPERATIONALIZING THE INDEPENDENT VARIABLES

The operationalization of most of the independent variables was very straightforward. Dummy variables distinguished likely voter samples from registered voter samples, tracking polls from standard polls, and weekday-only polls from weekday-and-weekend polls. Sample size was the number of respondents interviewed, and "days in the field" was the number of days interviewing was conducted; "day to election" was the number of days until November 3, averaged across every day the poll was in the field. Percent undecided represents a lack of crystallization in public opinion and simply equaled the average of 100 minus the sum of "true" support for Bush, Clinton, and Perot (as defined below) for every day a poll was in the field.

The only complicated independent variable is the measure of the strictness of defining a respondent as a supporter of a candidate. I created this variable by subtracting the sum of the "true" support for Bush, Clinton, and Perot (as defined below) averaged across every day a poll was in the field from the sum of the support for those three candidates as reported by the individual poll. Thus positive numbers suggest that a survey house was very strict in defining a respondent as a supporter of any of the three major candidates because that house was reporting fewer such supporters than in fact was the case, while negative numbers suggest lenient criteria. This difference was averaged together for every poll in my sample conducted by a specific survey house to create a summary measure for each of the 11 houses. Because the resulting scale is totally arbitrary, I recoded this variable to range between −.5 (the survey house with the most lenient criteria, on the average) and .5 (the survey house with the most strict criteria).

ESTIMATING "TRUE" SUPPORT FOR THE CANDIDATES

Any analysis of error rates assumes there is a valid criterion of "truth" or "reality" against which to judge estimation accuracy. For polls taken at the very end of the campaign, the criterion is obvious—the actual election results. This is the criterion utilized by both Crespi (1988) and Buchanan (1986) in their research. But we cannot use that same criterion for judging the accuracy of polls conducted earlier in the campaign unless we are willing to assume perfect stability in support for the candidate, at least during the last 2 months of the campaign. Despite the general success of the various "macro" models in predicting election outcomes from that far back in time,[8] this is not an assumption any of us would be willing to grant. Thus simply employing the actual election results as the criterion of accuracy will "build in" a large effect of date of survey (more specifically, days until the election) on accuracy, reducing the power of the analysis to detect other effects that might be correlated with the date of the survey.

A better method is to try to estimate changing levels of support for the candidates over time. One obvious solution is to combine the results of every available poll, weighting the results by sample size, to estimate support for the candidates on a daily basis. There were at least two polls in the field every day but two after September 1, and many times there were as many as four or five. The average number of respondents available for such a daily estimate was 1,363.

There are two problems with this solution, however. The first is aesthetic: the resulting average jumps around quite a bit from day to day, and I do not believe that public opinion during a presidential campaign is that labile. The more important problem is statistical: this procedure uses the results of a poll to help estimate the "reality" against which that poll's results are to be judged. This builds in a fair amount of statistical dependency, almost guaranteeing that polls with larger sample sizes, or survey houses that contributed more polls to the final sample—either of which would lead to a poll (or a survey house) contributing more to the estimate of reality—will be found to have smaller error rates.

I have come up with simple solutions to each of these problems. (1) To eliminate the dependency problem, the basic "reality" criterion was the average of all available poll results (weighted by sample size) except the poll whose accuracy is being judged. In other words, separate estimates of "truth" were devised for each polling organization based on the results of the other 10 polling houses. This results in each poll being judged against slightly different criteria, but it does eliminate the built-in dependency of utilizing all available data. (2) The daily estimates of support for each candidate are still somewhat choppy, however. To help "smooth out" the estimates I specified a time-series model predicting the daily average of these different estimates of support for each candidate (again weighted by sample size) from a lagged measure of support, plus dummy variables noting the major "events" of the last 2 months of the campaign (Perot's reentering the race, and the four debates that occurred in October).[9] The resulting "predicted" level of support for each candidate is, because of the nature of regression analysis, much smoother than the raw data on which it is based.[10]

---

[8] A general trend that in fact did not hold true in 1992: of the various popular models, roughly half picked Bush and half Clinton.

This method of operationalizing "actual" support for the candidates results in daily estimates. Those estimates are shown in figure 1. I simply averaged together these daily estimates for each day a poll was in the field to produce the criteria against which that poll was judged. To get an overall measure of "misestimation," I averaged the absolute value of the errors resulting from the estimation of support for each of the three candidates. This average misestimation of support per candidate is the dependent variable in the analyses presented in table 1.

Results

To establish the predictive validity of the measures of misestimation of public support for the candidates involved in the 1992 presidential election, I first examined the simple zero-order correlation between each predictor of misestimation rates and the criterion variable. These correlations also provide an initial examination of the eight hypotheses offered above, although more definitive tests are provided by the multivariate analyses to follow. For now, we will be most interested in the sign of the zero-order correlation, as predicted by the eight hypotheses.

The results of this preliminary analysis are shown in the first column

---

[9] Actual vote for each candidate was included as a final point to add a definitive note to the end of each time series. The effect of each debate was assumed to last 2 days.

[10] This procedure necessitated several fairly reasonable assumptions. For standard (i.e., not tracking) polls, I assumed that an equal number of interviews were conducted each day the poll was in the field, and that support for each candidate was constant during the interviewing. Obviously this assumption oversimplifies reality, but any errors it introduces into the analysis should be slight. Thus a poll's results were "counted" as one measure of actual public opinion for each day the poll was in the field, weighted by N/d, where N = the total sample size and d = the number of days of interviewing. It is possible to generate unique daily estimates from the tracking polls, however (i.e., the assumption of constant support over the period of interviewing is not required), because it is easy to calculate how much different any individual day's results must have been from the previous $d - 1$ days for the average level of support to change as much as it did from one day's average report (based on the previous d days of interviewing) to the next. Let me emphasize that this manipulation of the reported results of tracking polls was utilized only to increase the accuracy of my estimates of "reality." Tracking polls entered the basic data in exactly the same manner as all other polls—just as they were reported to the press.

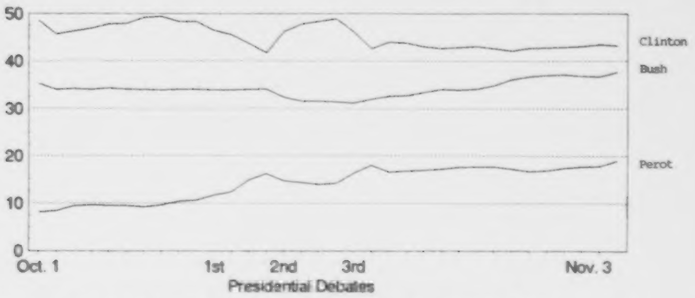


Figure 1. Estimated true public support for Bush, Clinton, and Perot during the last month of the 1992 campaign.

Table 1. Predicting Misestimation Rates in 1992 Trial Heat Polls

Predictor	Range	Zero-Order Correlation	Regression Coefficients
Sample size	575–2,086	−.10	.00005 (.00040)
Likely voters	0,1	.03	.31 (.31)
Number of days in field	1–5	−.25*	−.17* (.13)
Tracking poll	0,1	−.35***	−.59** (.26)
Weekdays only	0,1	.38***	.38* (.25)
Strictness of defining "supporter"	−.5–.5	−.09	.57 (.70)
Percent undecided	2.2–10.3	.13	.162 (.143)
Days to election	1.5–33	.04	−.016 (.033)
Survey house	0,1		−.73** (.33)
Constant			1.26** (.73)
R^2			.37
Adjusted R^2			.24
Standard error			.75

NOTE.—Regression coefficients are unstandardized *b* weights, with standard errors in parentheses. $N = 56$.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

of table 1. With only one exception, the results are as expected. Larger sample sizes are associated with smaller misestimation rates, as are a greater number of days in the field and a stricter definition by survey houses of "supporting" a candidate. Also as expected, polls conducted during the week are associated with greater error than polls that interviewed on weekends as well, as are a lack of crystallization in public opinion and a greater number of days until the election. Finally tracking polls, despite their negative reputation, are associated with lower misestimation rates. All of these relationships are consistent with hypotheses.

I also had hypothesized that results based on likely voters (as opposed to registered voters) would be associated with smaller error rates. Although this correlation is the smallest (in absolute magnitude) in table 1, it is positive, indicating that counter to expectations, likely-voter samples are associated with slightly higher error rates.

This last unexpected finding is one instance where, in retrospect, it would appear that politics were stronger than social science "theory." Algorithms designed to pick out likely voters that worked well in past elections may not have worked so well in 1992 (see Zukin, Carter, and Schulman 1987), with the added uncertainty from the off-again, on-again Perot candidacy and the effect this candidacy apparently had on electoral turnout. Perot received more votes (19 percent) than any third-party candidate had in over 70 years, and turnout, on a steady downward trend for almost 30 years, rebounded to levels that have not been seen since the 1960s.

MULTIVARIATE ANALYSES

The eight predictors all share some common variance, of course,[11] and the simple correlations in table 1 do not fully address the eight hypotheses listed above. To examine the independent effects of each of these predictors, I regressed the measure of misestimation rates on the eight predictor variables. After "forcing" these first eight predict[illegible] into the equations. I allowed dummy variables representing the different survey houses to enter the equation if the resulting coefficient was significant at the .10 level or better. One such dummy variable entered the equation presented in table 1, and one entered each of the equations presented in table 2. Because these "house" effects are of no theoretical interest in the absence of any good explanation for them, I make no further mention of them.[12]

---

[11] Indeed, in one case they share a lot of common variance: "percent undecided" and "days to election" correlate .93 with each other. Such a high correlation can inflate standard errors, reducing the power of the analysis to detect significant coefficients for the variables involved. If "days to election" is dropped from the equation, the coefficient for "percent undecided" is smaller than reported in table 1 (.097) but significantly different from zero ($p < .05$). Likewise, if "percent undecided" is dropped from the equation the coefficient for "days to election" switches to a marginally significant .018. Thus the predominant effect of these two variables on misestimation rates is positive, as expected, although when they are both in the equation neither variable alone explains significant amounts of unique variance. I decided to report analyses with both variables in the equation because the analyses in table 2 do not appear to suffer from inflated standard errors associated with these two variables.

The results are shown in the second column of table 1. Because I have theoretical expectations about the direction of every coefficient and because the sample size is relatively small, I adopt a one-tailed significance level. Of the eight variables of theoretical interest, three have important independent effects on error rates. The variable with the largest potential effect on misestimation rates is the number of days a poll was in the field. Each additional day reduced error rates by almost one-fifth of a point per candidate, or a total of one-half point for all three candidates. Thus 1-day polls included a full 2 percentage points more error than 5-day polls, according to my estimates.

The predictor with the strongest statistical effect is conducting a

[12] The "strictness of defining supporter" variable also represents house effects, but there we have a substantive interpretation of the variable. Only dummy variables representing survey houses that contributed at least five polls to my final sample were created. I should also note that I found no overall evidence that any of the 11 survey houses whose data I used in this article were any more or less accurate in their estimates than the others. A reviewer of this manuscript suggested that question wording might represent another important house effect that could be given a substantive interpretation. There were essentially three distinct question wordings utilized. Roughly a third of the surveys mentioned the party and the presidential candidates but made no mention of the vice presidential candidates. For example, Harris asked, "If the Presidential election were held today, who would you most likely vote for—George Bush for the Republicans or Bill Clinton for the Democrats or Ross Perot as an Independent?" Another third of the surveys mentioned the candidates' party affiliation before the presidential and vice-presidential candidates' names. For instance, Gallup asked, "If the presidential election were being held today, would you vote for the Republican candidates, George Bush and Dan Quayle, for the Democratic candidates, Bill Clinton and Al Gore, or for the independent candidates, Ross Perot and James Stockdale?" Finally, the remaining third of the surveys mentioned the presidential and vice-presidential candidates' names before their party affiliation. The CBS News/*New York Times* poll asked, "If the 1992 presidential election were held today, would you vote for George Bush and Dan Quayle, the Republican candidates, Bill Clinton and Al Gore, the Democratic candidates, or Ross Perot and James Stockdale, independent candidates for president?" I created dummy variables representing these three different question wordings and allowed them to enter the equations after the eight predictors of chief theoretical interest if they significantly improved the fit of the model. Only in the equation predicting overestimation of support for Bill Clinton (to be described later) did this question wording have any effect. Listing the candidates before their parties, or not mentioning the vice-presidential candidates at all, led to underestimation of the true support for Clinton relative to the question wording where the party is mentioned before the presidential and vice-presidential candidates. I have no substantive explanation for this single positive finding. I might note that several survey houses rotated the order in which the candidates were mentioned. As far as I could tell, there were no differences in results by order of listing the candidates.

tracking poll. As anticipated—and counter to their poor reputation—tracking polls are associated with almost six-tenths of a percentage point less error per candidate—or a total of well over 1.5 points for all three candidates—compared to standard polls. The coefficient associated with interviewing only during the week (and thus only in the evenings) is also quite strong. Polls conducted between Monday and Friday, with no interviewing on weekends, are associated with about four-tenths of a point more error per candidate—or a total of over 1 percentage point for the three candidates in the 1992 race.

Perhaps the most striking finding in these results, however, is the dog that did not bark: sample size, the variable that underlies sampling theory, the variable on which estimated "margins of error" are based, did not have a significant effect in the analyses. Neither the regression coefficient in the multivariate analysis nor even the zero-order correlation between sample size and misestimation rates came anywhere near statistical significance. This finding is very robust against alternative specifications of the multivariate model and various transformations of the variables involved in the analysis.[13] The case simply cannot be made with these data that a sample of 2,000 provided any more accurate results than a sample of 600.

Errors in Estimating Support for Individual Candidates

It is also instructive to see whether any of the variables in our analysis resulted in errors that over- or underestimated support for any of the individual candidates. This question is of more than academic interest; while I have focused so far on the overall accuracy of poll results—a focus that may be of most interest to survey researchers—questions about the direction of error for individual candidates are even more important because those errors could affect the outcome of the election itself. For instance, a close election can motivate more of the supporters of the candidates in the tight race to turn out and vote than would be the case if it looks like one candidate will be an easy winner (Ifill 1988). Potential supporters of a third-party candidate might switch to a major-party candidate if they think their first choice has no chance of winning but might stick with their candidate if it appears he or she might actually win (Black and Black 1993; see also West 1991).

---

[13] One of the reviewers of this manuscript suggested that sample size per se might not be related to misestimation rates, but that the reciprocal of the square root of sample size should be, as that is the format in which sample size enters the formula for calculating margins of error. I examined this and various other transformations of sample size but could not increase its correlation with misestimation rates to more than $r = .105$. For simplicity, I have retained the untransformed variable. This suggestion led me to examine the linearity of all of the zero-order correlations between the eight variables in table 1 and the measure of misestimation rates. These bivariate plots all led to a similar conclusion: leaving the eight independent variables untransformed but cubing the measure of misestimation rates provided the strongest zero-order correlations. When this single transformation is performed, the correlations in the first column of table 1 would read −.18, .00, −.27 ($p < .05$), −.28 ($p < .04$), .39 ($p < .01$), −.06, .28 ($p < .04$), and .18. None of the multivariate results change if the transformed dependent variable is employed, however, and again for simplicity I have utilized untransformed variables throughout.

Moreover, poll results could affect the actions of candidates themselves. In the past election, for example, an apparent surge in support for President Bush during the last few weeks of the campaign greatly energized his own stump speeches and general campaign efforts (Jehl 1992; Toner 1992). If this "surge" were mostly a methodological artifact based on a switch of sampling frames from registered voters to likely voters in tracking polls, then the methodological decisions being studied here could have real political effects (cf. Byrd 1992).

There are theoretical reasons to expect directional effects of several of these variables on reported levels of support for the specific candidates.

1. Traugott (1992) suggests that weekday-only polls produce a Republican bias, presumably because the middle-class voters who are the staple of the Republican party do not have to work evenings as frequently as working-class people, who are more likely to vote Democratic.

2. At least two studies have found that respondents reached on the initial call have a Democratic bias, compared with respondents reached on later callbacks (Carter, Shulman, and Zukin 1989; Traugott 1987). Thus I hypothesize that the number of days a poll is in the field—in this case, a surrogate for the number of callbacks possible—will be positively associated with overestimating support for Bush and negatively associated with overestimating suppprt for Clinton.

3. I would also speculate that the algorithms for selecting likely voters would have a Republican bias (because Republicans, in general, are more likely to vote than Democrats or Independents), and hence we might expect more overestimation of Bush support when the results are based on likely voters. If Perot voters are motivated by factors that are not part of the likely-voters algorithms, however, then we would also expect an underestimation of support for Perot when results are based on samples of likely voters.

4. Although most people were very familiar with President Bush during the 1992 campaign after almost 4 years as the chief executive (and many more on the national stage), Bill Clinton and Ross Perot were relatively new political figures. Thus if large numbers of the pub-

lic are weak in the strength of their support for their favorite candidate, it is because they are uncertain about Clinton and Perot, not because they are uncertain about Bush. Hence I would hypothesize that survey houses that adopt relatively strict definitions of "supporting" a candidate would tend to underestimate support for Clinton and Perot much more than for Bush.

I calculated measures of overestimation of support for each candidate by subtracting the "actual" levels of support for each candidate (as determined by the method described above) during the days a survey was in the field from the level of support reported by that poll. Positive numbers on these difference scores mean a poll overestimated support for the candidate (relative to my estimate of reality), while negative numbers suggest the poll underestimated the actual level of the candidate's public support. These measures were then regressed on the same set of predictors employed in the analyses in table 1.

The results are shown in table 2. Traugott suggested that polls conducted on weekdays only would overestimate support for Bush and concomitantly tend to underestimate support for Clinton. There is some support for this hypothesis, in that polls conducted exclusively during the week overestimate support for Bush by a full percentage point, relative to polls in which some of the interviewing was on weekends. This predictor has the expected negative sign for Clinton but is not reliably different from zero.

There is strong support in the data for the hypothesis that likely-voter samples would have a pro-Bush and anti-Perot bias compared to registered-voter samples. The effect of sampling frame is almost 2 points for each candidate, resulting in an average overestimation of the apparent lead Bush had over Perot of more than 4 percentage points.

There is also strong support in the data for the hypothesis that stricter operationalizations of "supporting" a candidate would hurt apparent levels of support for the two newcomers, Clinton and Perot, more than Bush. The effect is highly significant for Clinton and marginally significant for Perot, while it is positive (albeit not reliably different from zero) for Bush. This variable alone contributed almost 3.5 points to the total underestimation of support for Clinton in the data, and almost a full point to the underestimation of support for Perot.

There is no support for the hypothesis that number of days in the field would be associated with preference for the Republican candidate. However, the data do show three other results that were not anticipated. First, the greater the number of days until the election, the more support for Bush is overestimated, and the more support for Perot is underestimated. The magnitude of this error is about a 4.5-percentage-point greater overestimation of support for Bush on

Table 2. Overestimation of Support for Each Candidate

	Bush		Clinton		Perot	
Sample size	.0006	(.0009)	.0015*	(.0009)	−.0018*	(.0010)
Likely voters	1.95***	(.59)	−.45	(.60)	−2.21***	(.68)
Days in field	.08	(.29)	.13	(.30)	.14	(.34)
Tracking poll	−.11	(.59)	1.10*	(.61)	−.18	(.70)
Weekdays only	1.00**	(.55)	−.30	(.54)	−.51	(.62)
Strictness of defining "supporter"	1.10	(1.77)	−9.21***	(1.56)	−2.51*	(1.79)
Percent undecided	−.27	(.32)	.25	(.32)	1.11***	(.36)
Days to election	.14*	(.07)	−.02	(.07)	−.27***	(.08)
Survey house	1.38*	(.82)	2.75***	(.94)	−2.21**	(1.08)
Constant	−2.70*	(1.63)	−4.50***	(1.78)	−.30	(2.04)
R^2	.35		.52		.39	
Adjusted R^2	.23		.43		.27	
Standard error	1.68		1.66		1.90	

NOTE.—Table entries are unstandardized regression coefficients, with standard errors in parentheses. $N = 56$.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

October 1 relative to November 2, with an underestimation of support for Perot of almost twice that magnitude over the same period. It is possible that the survey houses were making subtle, unreported methodological decisions that hurt the apparent support for Perot until the very end of the campaign, when they were convinced it was real.

A second unexpected result is the effect of sample size on support for Clinton and Perot. Larger samples are associated with overestimation of support for Clinton and underestimation of support for Perot. A final surprising finding is the effect of opinion crystallization on estimates of support for Perot. The less crystallization of opinion in the public as a whole, the greater the overestimate of support for Perot. Notice that this effect is independent of and opposite in direction from the strictness of a survey house's definition of a candidate's "supporters." I have no good explanation for either of these last two findings.

Discussion

The research process is always a trade-off between what is best from a purely scientific perspective, and what is possible within the constraints of available resources. The tensions in this trade-off are exaggerated when the basic undertaking—a nationally representative sample survey—is very expensive to begin with. This study is an attempt to provide some solid empirical evidence for the consequences of different methodological decisions that are regularly made in the process of trying to conduct the best possible survey for a given budget, so that those methodological decisions will be informed ones.

The most important variables by far in predicting overall error rates are the number of days a poll is in the field, and the decision to conduct a survey between Monday and Friday, with no interviewing conducted on the weekends. A 1-day poll conducted during the week resulted in over a full percentage point more error in estimating support for a candidate than would be expected from a 5-day poll with interviewing conducted on the weekends and during the week. Thus these two variables contributed 3 full percentage points of error in the three-candidate race of the 1992 presidential election.

Turning to factors that predicted over- or underestimation of support for the specific candidates, two stand out as particularly noteworthy, and likely to replicate across different elections and different sets of specific candidates. First, when interest in a campaign (and subsequent turnout) is unusually high, sample frames of likely voters will tend to overestimate support for Republican candidates and underestimate support for any popular third-party candidates, relative to registered-voter sampling frames. Second, survey houses that follow strict criteria for defining a respondent as "supporting" a candidate will tend to underestimate support for less well known candidates relative to sitting incumbents.

In many ways this study complements a previous, and more elaborate, study conducted by Crespi (1988). Like him, I found that sample size did not significantly affect error rates in trial heat polls, once other variables were considered in a multivariate analysis. In certain ways his findings are more generalizable than mine, because he analyzed trial heat polls conducted over several different election years for offices at national, state, and local levels, while my data are restricted to national samples targeting the presidential election in a single election year. But my analysis includes more variables endogenous to the survey research process and thus can provide more guidance to researchers allocating scarce resources toward improving data quality.

There are other ways to improve the accuracy of trial heat polls (or any survey, for that matter) other than varying the factors included in this study, of course. Of the six variables considered here over which survey researchers have some control, three—the number of days a survey is in the field, interviewing only on weekdays as opposed to weekend interviewing as well, and conducting a tracking poll with a "rolling" sample design—could all influence response rates. Increasing the number of callbacks of every unsuccessfully contacted number before that number is abandoned, and/or greater efforts to "convert" initial refusals into respondents, may be more cost-effective ways of improving response rates, and thus survey accuracy. In a way it is quite distressing that the data show that tracking polls result in less error when we know all of the methodological compromises they entail, because it suggests how badly we are doing in obtaining the most accurate possible results in the rest of our surveys. At the very least, I would call for the regular reporting of response rates along with the major "results" of surveys so that this information would become part of the public record, and thus directly available for refining future research on survey accuracy.

It is also distressing that sample size--the one variable that goes into calculations of "margins of error" that have become standard practice to report—was not related to misestimation rates. Like the drunk searching for his lost car keys under a street lamp simply because he can see better there, we may be fooling ourselves into thinking that reports of margins of error based solely on sample size are a good measure of survey misestimation rates just because the mathematics is very tractable.

If my results were the only ones to report a null finding for sample size, I would proceed much more cautiously. But they are not. My results echo a previous null finding reported by Crespi (1988), who had completely different data, a somewhat different model, and a very different measure of true public opinion. These two studies together lead me to recommend that the common practice of reporting margins of error based solely on sample size be abandoned because it gives us a false sense of security. A new and more empirically defensible standard must be developed by the field, one based more on response rates and some estimate of the nature of the public opinion being measured than on sample sizes alone.

References

Black, G. S., and B. D. Black. 1993. "'Perot Wins!' The Election That Could Have Been." *Public Perspective* 4(2):15–16.

Bradburn, N. M. 1992. "Presidential Address: A Response to the Nonresponse Problem." *Public Opinion Quarterly* 56:391–97.

Brennan, J. 1992. "In Polls, 1 Plus 1 Doesn't Always Equal 2." *Los Angeles Times*, October 29.

Broder, D. S. 1988. "Did I Say That?" *Washington Post*, December 28.

Buchanan, W. 1986. "Election Predictions: An Empirical Assessment." *Public Opinion Quarterly* 40:222–27.

Byrd, J. 1992. "Poll Play." *Washington Post*, October 4.

Carter, B., M. A. Schulman, and C. Zukin. 1989. "The Impact of Day-of-the-Week, Time-of-Day, and Callbacks on Candidate Preference in Pre-election Surveys." Paper presented at the forty-fourth annual meeting of the American Association for Public Opinion Research, St. Petersburg Beach, FL.

Converse, P. E., and M. W. Traugott. 1986. "Assessing the Accuracy of Polls and Surveys." *Science* 234:1094–98.

Costikyan, E., and L. A. Kaplan. 1984. "Why Trust the Polls?" *New York Times*, November 5.

Crespi, I. 1988. *Pre-election Polling: Sources of Accuracy and Error*. New York: Russell Sage.

Crespi, I., and D. Morris. 1984. "Question Order Effect and the Measurement of Candidate Preference in the 1992 Connecticut Elections." *Public Opinion Quarterly* 48:578–91.

Houston, P. 1992. "Clinton's Surge Largest, Fastest in History of Polling." *Los Angeles Times*, August 7.

Hugick, L., G. Molyneux, and J. Norman. 1993. "The Performance of the Gallup Tracking Poll: The Myth and the Reality." *Public Perspective* 4(2):12–14.

Ifill, G. 1988. "Voter Abstention Breaks 1948 Record: Experts Cite Negative Campaigning, Polls." *Washington Post*, November 10.

Jehl, D. 1992. "'Fired Up' Bush Heads West, Blasts Clinton on Trust Issue." *Los Angeles Times*, October 26.

Morin, R. 1989. "We Pollsters Had It Coming." *Washington Post*, November 13.

———. 1992. "Surviving the Ups and Downs of Election '92." *Washington Post Weekly Edition* 10(2):37.

Public Perspective. 1992a. "The Polls in 1992: A Roper Review." *Public Perspective* 4(1):11.

———. 1992b. "Trial Heats: Nationally . . ." *Public Perspective* 4(1):100–101.

———. 1993. "Problems in Exit Polling: Interviews with Warren J. Mitofsky and John Brennan." *Public Perspective* 4(2):19–23.

Richburg, K. B. 1984. "Every Poll Had Right Winner; Their Margins Varied Widely." *Washington Post*, November 8.

Roper, B. 1983. "Election Poll Errors: A Problem for Us All."*AAPOR News* 10(2):1.

Rosenstiel, T. B. 1992. "Pollsters Search for Reasons for Disparities in Surveys." *Los Angeles Times*, October 27.

Taylor, H., and D. Krane. 1993. "The Polls and the 1992 Elections: Lessons Learned." *Public Perspective* 4(2):10–12.

Toner, R. 1992. "Contest Tightens as Perot Resurges and Clinton Slips: Polls Buoy Bush's Camp." *New York Times*, October 25.

Traugott, M. W. 1987. "The Importance of Persistence in Respondent Selection for Preelection Surveys." *Public Opinion Quarterly* 51:48–57.

———. 1992. "A Generally Good Showing, but Much Work Needs to Be Done." *Public Perspective* 4(1):14–16.

West, D. M. 1991. "Polling Effects in Election Campaigns." *Political Behavior* 13:151–64.

Zukin, C., B. Carter, and M. A. Schulman. 1987. "Trade-offs in Respondent Selection Methods: Theoretical Purity versus Real World Concerns." Paper presented at the forty-second annual meeting of the American Asociation for Public Opinion Research, Hershey, PA.

QUESTION WORDING AND PARTISANSHIP

CHANGE AND CONTINUITY IN PARTY LOYALTIES DURING THE 1992 ELECTION CAMPAIGN

PAUL R. ABRAMSON
CHARLES W. OSTROM

Abstract Gallup macropartisanship varies more over time than aggregate measures of partisanship employing the standard Michigan Survey Research Center (SRC) party identification measure, but previous analyses do not provide direct evidence about why Gallup macropartisanship is more variable. Although these differences could result from the short-term focus of the Gallup party affiliation question, aggregate-level analyses cannot test the effects of question wording directly. Between March and October 1992, we conducted a series of question-wording experiments, employing six statewide computer-assisted telephone interview (CATI) surveys of Michigan adults, including a four-wave panel study. Our analyses strongly suggest that the Gallup measure responds more to short-term political conditions and clearly demonstrate that the Gallup measure is less stable over time. These individual-level results help explain why Gallup macropartisanship varies more over time than aggregate measures of partisanship employing the standard SRC measure and raise questions about the degree to which one can generalize from analyses using the Gallup data to the research literature on party identification.

PAUL R. ABRAMSON and CHARLES W. OSTROM are professors of political science at Michigan State University. Their research was supported by an All-University Research Initiation Grant from Michigan State University and by funding from the Institute for Public Policy and Social Research, the College of Social Science, the Office of the Provost, the Office for Research and Development, and the Office of the Vice-Provost for University Outreach, Michigan State University. The authors are grateful to Michael J. Keefe and Sara McLaughlin for their assistance with the data analysis and to Karen Clark for her assistance in designing the September and October surveys. John H. Aldrich, George F. Bishop, William Claggett, Robert S. Erikson, Ada W. Finifter, Jim Granato, Donald Philip Green, Larry Heimann, Michael B. MacKuen, Sara McLaughlin, Warren E. Miller, and Brian D. Silver made helpful comments on earlier versions of this article, and Gregory B. Markus made helpful suggestions.

Public Opinion Quarterly Volume 58:21–48 © 1994 by the American Association for Public Opinion Resear
All rights reserved. 0033-362X/94/5801-0001$02.50

Party identification is the most widely used attitudinal variable in the study of American politics. However, during the past 2 decades scholars have disagreed about the origins of partisanship, its stability among individuals and electorates, the best way to measure partisan loyalties, the extent to which party identification shapes or is shaped by policy preferences, the dimensionality of the party identification measure, and the meaning of partisan independence. In recent years, controversies about the stability of party identification have sharpened. Whether party identification is stable or malleable is central to theories of electoral behavior. In their classic study *The American Voter*, Angus Campbell and his colleagues (1960) viewed party identification as a basic and highly stable social-psychological attachment that structured the way Americans evaluated candidates and viewed political issues. In contrast, revisionist theories maintain that party identification can be highly reactive to current evaluations of economic and political conditions (see, especially, Fiorina 1981).

Two basic partisanship questions have been employed in this controversy. The most widely used question is the standard party identification question developed by the University of Michigan Survey Research Center (SRC), which is designed to measure long-term partisan attachments. That question reads as follows: "Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?"[1] The SRC question has been employed biennially by the National Election Studies (NES) since 1952 and has been employed annually by the General Social Surveys (GSS) since 1972.

The Gallup organization has also measured partisan loyalties, using a measure it labels "party affiliation." The Gallup question reads as follows: "In politics, as of today, do you consider yourself a Republican, a Democrat, or an Independent?" The Gallup question is included in most Gallup surveys and is therefore available to researchers on a monthly or quarterly basis.

Controversies about the individual-level stability of partisanship have relied largely upon panel analyses of the SRC measure (see Achen 1975; Alwin, Cohen, and Newcomb 1991; Alwin and Krosnick 1991; Asher 1974; Brody 1991; Brody and Rothenberg 1988; Fiorina 1981; Franklin 1992; Green and Palmquist 1990, 1992; Hunter and Coggin 1976; Lockerbie 1989; Miller 1992; Niemi and Jennings 1991). Donald Green and Bradley Palmquist (1990, 1992) present the most compelling evidence to date using these data. Their findings suggest that, once measurement error is taken into account, partisan loyalties, as measured by the Michigan SRC party identification question, are highly stable over time.

---

[1] The Michigan SRC also employs follow-up questions that measure the strength of partisanship among Republicans and Democrats and that determine whether self-proclaimed Independents feel closer to the Republican or Democratic party. We employ these follow-up questions to develop an alternative measure of partisanship that includes Independent leaners as partisans.

In recent years there have also been several time series analyses of aggregate levels of partisanship among the electorate (see Abramson and Ostrom 1991, 1992; Allsop and Weisberg 1988; MacKuen, Erikson, and Stimson 1989, 1992a, 1992b; Stimson 1991; Weisberg and Smith 1991). These studies utilize several partisanship questions but rely mainly on the Gallup party affiliation question because the Gallup data provide a finer grade measure of partisan loyalties over time. Michael MacKuen, Robert Erikson, and James Stimson (1989) provide the most significant analyses of the Gallup data by studying change in Gallup partisanship from 1945 through 1987. They find substantial changes in Gallup macropartisanship (i.e., party affiliation at the aggregate level). They suggest that changes in macropartisanship follow a structured pattern within presidential administrations, that they are related to short-term electoral outcomes, and that they are driven strongly by short-term economic and political evaluations. On the basis of these results, they argue that the traditional view of party identification must be substantially reformulated.

In a recent study (Abramson and Ostrom 1991), we use NES, GSS, and Gallup surveys to study aggregate-level changes in partisanship, analyzing both the SRC measure and the Gallup measure taken at 1- and 2-year intervals. While we find that both measures of macropartisanship change over time, Gallup macropartisanship varies more than does SRC macropartisanship. As we show, compared with the Gallup results, the NES and GSS results reveal less short-term variation and display less total variation. The NES partisanship results are less related to short-term electoral outcomes and do not appear to be driven strongly by short-term economic and political evaluations. Even though our analyses clearly demonstrate that the SRC measure responds to short-term economic evaluations, our results are more compatible with *The American Voter* formulation (Campbell et al. 1960).

We account for these differences by concluding that the Gallup party affiliation question taps a different aspect of partisanship than the SRC question. On their faces, the SRC ("generally speaking" . . . "usually") and the Gallup ("as of today") items appear to have a different time referent, with the Gallup item having a short-term focus. Indeed, Converse (1976) made this point 15 years earlier, and Converse and Pierce (1986) argue that partisanship questions lacking a long-term component measure a fundamentally different phenomenon than the standard SRC question. As they write: "Among other wording requirements, experience has shown that if the notion of custom or generality is not conveyed, as by 'habituellement' (usually), the responses reasonably enough lose their long-term meaning. When the phrase 'generally speaking' is dropped from the parent United States question, for example, the responses come to behave over time very much as statements of momentary vote preference, thereby obviating the whole point of the party identification measurement" (Converse and Pierce 1986, p. 73). If Converse and Pierce are correct, the Gallup measure should be substantially more responsive to short-term political events and conditions than the Michigan SRC measure is.[2] If this line of argument is valid, the MacKuen, Erikson, and Stimson findings (1989, 1992a) may be an artifact of using the Gallup question.

The contradictory findings generated from analyses of the Gallup and SRC measure are difficult to reconcile with one another. MacKuen, Erikson, and Stimson argue (1992b, p. 479) that "the important matter of reconciling macro movement of partisanship with micro theory about individual partisanship . . . remains a challenge." But if differences between the Gallup and SRC results are an artifact of question wording, complex efforts to resolve this problem are not necessary. Social scientists are now well aware that even relatively small differences in question wording can lead to substantially different results (see Schuman and Presser 1981). Even so, with the exception of two papers by George Bishop, Alfred Tuchfarber, and Andrew Smith (1992, 1993), none of the evidence marshaled to date tests the effects of question wording directly. Differences between aggregate-level Gallup and SRC findings could result from question wording differences, but they could also result from the differential timing of polls that use the Gallup and SRC questions, from different procedures used by these polling agencies, or from the different context in which the SRC and Gallup questions are asked (see Bishop, Tuchfarber, and Smith 1992, 1993).

To determine whether question wording differences per se lead to different findings, we conducted a series of question-wording experiments. Our results suggest that responses to the Gallup question are more related to short-term political evaluations than are responses to the SRC question and that responses to the Gallup question are less stable over time. Our individual-level results help explain why Gallup macropartisanship varies more over time than do aggregate measures of partisanship employing the SRC measure and raise questions about the extent to which one can generalize from findings based on Gallup data to the research literature on party identification.

---

[2] Analyses by Borrelli, Lockerbie, and Niemi (1987) suggest that the Gallup measure might have systematic biases that favor the party advantaged in the most recent presidential election.

Question-wording Experiments

Our question-wording experiments were conducted as part of a series of computer-assisted telephone interview (CATI) statewide surveys measuring public policy preferences in Michigan during the 1992 election campaign. The surveys are briefly described in table 1. As can be seen, we conducted a four-wave panel study in which the question formats were varied randomly among respondents. In addition, we conducted two cross-sectional studies in which all respondents were asked both the Gallup and SRC questions. A detailed discussion of the sampling procedures is provided in the Appendix.

In addition to the partisanship questions, all six surveys asked questions designed to measure short-term evaluations and electoral preferences. All six measured presidential approval (using the standard Gallup question), gubernatorial approval, and presidential vote preference. All but the first survey included the five items employed to create the index of consumer sentiment (using the standard Michigan SRC questions).[3] In addition, the final three surveys asked respondents to evaluate George Bush's handling of the economy and his handling of "our relations with foreign countries."

Because of the random allocation of the partisanship questions, there is no reason to expect any systematic differences among subsets of respondents who are asked different questions. In fact, there is very little difference among these subsets.[4] Thus, if relationships differ among respondents who are asked different partisanship questions, they are very likely to result from differences in question wording.

The Political Context

When short-term forces are weak, the Gallup and SRC questions are likely to yield similar results—there will be low levels of variability. When the political context is volatile, both measures will change, but the degree of change will differ. If Converse and Pierce (1986) and if we are correct, the preface to the Gallup question will amplify the

---

3. The items used to construct this index are reported in Curtin (1982). We constructed an additive index (using listwise deletion), with scores ranging from a low of 0 to a high of 10.

4. Panel survivors differ somewhat from those who were not reinterviewed. For example, they are somewhat more likely to be Republicans, to approve of Bush's job performance, to be strong party identifiers, and to be white. But these differences are small, and there is no evidence of differential attrition among panelists who are asked the different question formats.

Table 1. Summary of Question-wording Experiments

Survey Dates (1992)	Type of Sample	Number of Respondents	Partisanship Questions
March 8–12	First wave of panel survey	1,066	Half asked the Michigan SRC questions; half asked the Gallup question. The question format was randomly assigned.
April 12–24	Cross-sectional	997	All asked both the SRC questions and the Gallup question. Half asked the SRC questions at the beginning of the interview and the Gallup question at the end; half asked the Gallup question at the beginning of the interview and the SRC questions at the end. The question order was randomly assigned.
May 14–22	Second wave of panel survey	836	Half the respondents who were asked the SRC questions in the first wave were asked the SRC questions again, half were asked the Gallup question. Half the respondents who were asked the Gallup question in the first wave were asked the Gallup question again, half were asked the SRC questions. The question format was randomly assigned.
July 19–August 4	Cross-sectional	886	Same as the April 12–24 cross-sectional survey.
September 12–19	Third wave of panel survey	772	Respondents asked the SRC questions in the first wave were asked the SRC questions again; respondents asked the Gallup question in the first wave were asked the Gallup question again.
October 20–26	Fourth wave of panel survey	737	Respondents asked the SRC questions in the second wave were asked the SRC questions again; respondents asked the Gallup question in the second wave were asked the Gallup question again.

Source.—The surveys were computer-assisted telephone interview (CATI) statewide surveys of Michigan residents 18 years old and above. They were conducted by the Survey Research Division of the Institute for Public Policy and Social Research, Michigan State University. The six questionnaires used in our surveys are available on request.

impact of short-term forces, whereas the preface to the SRC question will attenuate their impact.

Our surveys began shortly before Super Tuesday (March 10, 1992), when many pundits expected an unpopular Bush to be reelected. By the third survey H. Ross Perot had begun to lead Bush in national public opinion polls, as well as in polls in several key states. By the July–August survey, Perot had withdrawn from the presidential contest (on July 16), and Bill Clinton held a strong lead over Bush in national opinion polls, as well as in many statewide polls. By our September survey, Clinton's lead over Bush had narrowed somewhat, and Perot had not yet reentered the presidential race. By our final survey, Perot had reentered the contest (on October 1) and had participated in the three presidential debates (October 11, 15, and 19). Bush was trailing badly in the polls, and Perot was surging. Most Michigan polls, including ours, showed Clinton with a 10–15-percentage-point lead over Bush.

During these 7 months the presidential contest changed dramatically in other ways. During our first survey, the Democratic nomination contest was still competitive, but by May Clinton was the presumptive nominee. For our final three polls, Clinton was the official Democratic nominee. While Bush was clearly the presumptive Republican nominee in March, Pat Buchanan was still campaigning actively. By May, Bush had won enough delegates to assure his renomination, although approval for his presidential performance continued to erode. Our July–August poll began 3 days after the Democratic convention (July 13–16), and ended 13 days before the Republican convention began (August 17–20). By September, Bush was the official Republican nominee.

Political competition varied from state to state. Michigan is similar to the United States overall, demographically, although its electorate is marginally younger than the U.S. electorate and has a slightly larger percentage of African-Americans. It is a competitive two-party state, although it has historically been somewhat more Republican than most non-Southern states. In 1992, Michigan voted 44 percent for Clinton, 36 percent for Bush, and 19 percent for Perot (compared with 43 percent, 37 percent, and 19 percent, respectively, for the nation as a whole).

Although the political climate in Michigan was similar to that of the nation as a whole, it did have some special characteristics. Michigan was holding its first presidential primary since 1980, and the first since 1976 that was officially recognized by the Democratic National Committee. Over the course of the campaign, Michigan became a "battleground" state. Bush and Clinton both made numerous visits, and Michigan was the site of the final presidential debate (October 19). The hotly contested presidential race may have politicized the Michigan electorate. Despite the absence of any major statewide races (John Engler, the Republican governor, was in the second year of his 4-year term, and neither of the state's two Democratic Senators was up for reelection), turnout rose seven points, compared with five points for the nation as a whole. Michigan voters, along with voters in 13 other states, approved a ballot proposal limiting congressional terms. But all 11 House incumbents who ran in the general election were reelected. Michigan elected 10 Democrats and six Republicans to the 103d Congress, close to the 258-to-176 ratio for all major-party House seats.

In short, Michigan may be distinctive, but it is scarcely atypical. If we find that the Michigan electorate reacts differently to the Gallup and SRC partisanship questions, it is highly likely that similar results would be found among the national electorate.

Aggregate Differences

Aggregate differences in partisanship must be assessed in light of the prevailing political context. Overall levels of presidential approval provide the single most visible quantitative indicator of the political context. High levels of approval are usually associated with peace, economic prosperity, and the absence of scandal, although presidents often experience temporary surges of approval in periods of international crisis. When assessing presidential approval, the 50 percent benchmark takes on special significance. From 1956 through 1992, seven presidents have sought to be returned to the White House. All four with Gallup approval ratings above 50 percent were successful; all three with ratings below 50 percent were defeated. When presidential approval is low, the political context often becomes quite volatile. In March 1992, when our surveys began, Gallup polls revealed that only 41 percent of the American public approved of Bush's performance as president, and by the end of July his approval rating fell to 29 percent. Although his approval rating improved somewhat, no Gallup poll between then and the election registered an approval rating above 40 percent.

Based on previous research (see Abramson, Aldrich, and Rohde 1991; Abramson and Ostrom 1991; MacKuen, Stimson, and Erikson 1989), we expect both the Gallup and the SRC measure to be responsive to presidential approval. But, if we are correct, the Gallup measure will be more responsive than the SRC measure. Furthermore, given that Bush is a Republican, the impact of the political context will affect Republicans and Democrats differently. Our thesis suggests that the Republican party will register higher levels of support with the SRC measure than with the Gallup measure. Disapproval with Bush will be less likely to affect overall levels of support for the Democratic party. Democrats will remain Democrats, whereas Republicans disaffected with Bush may become Independents. But Republicans will not necessarily switch their loyalties to the Democratic party, and Independents disaffected with Bush are likely to remain Independents. Therefore, support for the Democrats should be similar regardless of which measure is used. Our thesis was essentially silent about the proportion of Independents. In fact, NES surveys have consistently found more Independents than have the Gallup surveys (see Abramson 1983). Even so, if the Gallup question is more responsive to short-term political conditions one might expect more Independents with the Gallup measure. Many voters may have been rethinking their partisanship, and there was a strong Independent presidential candidate for much of this period.

Table 2 presents the distribution of partisan loyalties in all six surveys, controlling for question type. As the table indicates, all three hypotheses are supported. More Republicans are found in all six surveys when the SRC question is used, with differences ranging from 1.1 to 6.0 percent. Differences in Democratic support are quite small, the largest being only 1.3 percentage points. And in five of the six surveys, there are more Independents when the Gallup question is used, with differences ranging from 2.5 to 6.8 percent. Although differing levels of Republican support and of Independence are not significantly different at the conventional .05 level, the directional nature of all three hypotheses is supported.[5]

Table 2 also presents the MacKuen, Erikson, and Stimson macropartisanship score. Their measure is simply the percentage of partisans who say they are Democrats. Together, the three separate hypotheses predict that Gallup macropartisanship should be more Democratic than SRC macropartisanship. As table 2 reveals, in five of our six surveys, the Michigan electorate is more Democratic with the Gallup measure. In March, May, and October, the Gallup question reveals that a slight majority of partisans are Democrats, whereas the Michigan SRC measure reveals a slight Republican majority.[6] Throughout this period in

---

[5] When we compare the distribution of Republicans, Independents, and Democrats, we found that the distribution was significantly different at the .05 level (2 *df*) in the May survey. The differences for the March survey fell just short of significance at the conventional .05 level ($p = .06$).

[6] For both our April and July–August surveys we compared the aggregate distribution of partisanship among respondents who were asked the SRC questions or the Gallup question early in the interview and those who were asked the SRC questions or the Gallup question late in the interview. For the April survey, there were virtually no differences between the distribution of responses among respondents asked these questions early and late in the interview. For the July–August survey, we found that respon-

Table 2. Distribution of Partisanship, by Question Format

	Survey Dates (1992)											
	March 8–12		April 12–24		May 14–22		July 19–August 4		September 12–19		October 20–26	
Partisanship	Michigan SRC (%)	Gallup (%)	Michigan SRC (%)	Gallup (%)	Michigan SRC (%)	Gallup (%)	Michigan SRC (%)	Gallup (%)	Michigan SRC (%)	Gallup (%)	Michigan SRC (%)	Gallup (%)
Republican	31.4	25.4	28.0	26.9	33.3	26.7	27.5	26.4	31.0	28.2	31.4	27.7
Independent	37.3	44.1	41.8	44.3	38.2	45.8	39.5	38.6	40.3	42.8	36.5	39.6
Democrat	27.8	26.7	25.1	24.0	27.4	27.2	30.0	31.2	26.9	27.9	30.3	31.6
Other response	.9	1.0	2.9	.8	1.2	.3	.3	.0	.3	.3	.3	.0
Don't know	2.7	2.9	2.3	4.0	.0	.0	2.6	3.9	1.5	.8	1.6	1.1
Total percent	100.1	100.1	100.1	100.0	100.1	100.0	99.9	101.0	100.0	100.0	100.1	100.0
Number	526	524	970	974	427	389	869	869	390	376	370	361
Macropartisanship score[a]	47.0	51.3	47.2	47.2	45.2	50.5	52.2	54.2	46.5	49.8	49.1	53.3

Source.—See table 1.

[a] Percentage of major-party supporters who are Democrats.

which Bush had low approval ratings, the ratio of Democrats to Republicans was higher with the Gallup measure. The consistency of these results strongly suggests that the Gallup measure is more responsive to volatility in the political context.[7]

Short-Term Evaluations and Electoral Preferences

Our previous aggregate-level analyses (Abramson and Ostrom 1991, tables 3 and 4) demonstrate that changes in Gallup partisanship are more strongly related to short-term evaluations and electoral outcomes than are changes in partisanship as measured by the standard Michigan SRC questions. Because the Gallup surveys seldom include questions that allow researchers to identify Independents who lean toward a major party, we used a single measure of Gallup macropartisanship—the percentage of partisans who said they were Democrats. On the other hand, we constructed two measures of Michigan SRC partisanship. The first, which we labeled NES$_1$ (or GSS$_1$) was the percentage of party identifiers who said they were Democrats; the second, which we labeled NES$_2$ (or GSS$_2$) was the percentage of Democratic identifiers and Independents who "leaned" toward the Democratic party divided by the total percentage of party identifiers and Independent leaners. In most cases, aggregate change in the second measure was more strongly related to short-term evaluations and electoral outcomes (see Abramson and Ostrom 1991, tables 3 and 4). At the individual level, however, including leaners might lead to weaker relationships. Relationships between short-term political preferences and partisanship are always highest among strong party identifiers, and when Independent

---

dents who were randomly chosen to be asked the SRC questions early and the Gallup question late were marginally more Democratic than those who were randomly chosen for the alternative format.

[7] Bishop, Tuchfarber, and Smith (1992, 1993) present comparisons of the aggregate differences in partisanship in 15 telephone polls that they conducted in Ohio between April 1991 and June 1993. Only one of their surveys found a significant difference between results using the Gallup measure and those using the Michigan SRC format. They conclude that the greater volatility of Gallup partisanship during the postwar years is unlikely to result from differences in question wording. In fact, they do find some interesting differences, although these differences fell short of statistical significance. Their earliest study was a statewide survey in Ohio conducted in April 1991. This is an important period, since several analysts have pointed to the surge in Republicanism in national polls using the Gallup measure conducted shortly after the Persian Gulf War (see Abramson and Ostrom 1992; Kagay 1991; Kohut 1991). Among Ohio respondents who were asked the Gallup question, 50.8 percent of those who supported a major party (N = 181) were Republicans; among respondents who were asked the SRC question, 42.8 percent of those who identified with a major party (N = 201) identified as Republicans.

leaners are classified as partisans, the proportion of partisans who are strong party identifiers is reduced.

For table 3 we construct the microlevel equivalents of our measures and show their relationship to short-term evaluations and electoral preferences. The first measure, SRC_1, is based on party identifiers and excludes all Independents; SRC_2 includes Independents who lean toward either the Republican or Democratic party.[8] For the Gallup question we construct a single measure that excludes Independents. All correlations are calculated so that the expected zero-order relationships are positive. Our thesis predicts that short-term evaluations and voting preferences will be more strongly related to the Gallup measure.[9] For every comparison, the correlation between Gallup partisanship and the short-term indicator should be higher than the correlation between the SRC measure and that indicator. The relationship should hold regardless of whether we employ the SRC measure based only on self-proclaimed partisans (SRC_1) or the SRC measure that classifies Independent leaners as partisans (SRC_2).[10]

Most of our results support our basic thesis. In all six surveys, Gallup partisanship is more strongly related to the standard measure of presidential approval than either of the two SRC measures. In all three surveys rating approval of Bush's economic performance, Gallup partisanship is more strongly related to approval than are either of the two SRC measures. In all three of these surveys, Gallup partisanship is more strongly related to approval of Bush's foreign policy performance than either of the two SRC measures. In five of the six surveys, Gallup partisanship is more strongly related to gubernatorial approval than either of the SRC measures.

Because of the changing nature of the presidential race, we altered our measure of presidential vote preference across the six surveys. In March, we paired Bush against "any Democrat," whereas in April we

---

[8] We use the acronym SRC rather than NES to emphasize that these surveys employ the Michigan SRC measure but were not conducted as part of the National Election Studies.

[9] Analyses by Ostrom and Simon (1988) of the 1980 NES yearlong panel study strongly suggest that presidential approval is more responsive to short-term fluctuations than is the Michigan SRC measure of party identification. Using our own surveys, we measured the over-time stability of presidential approval, gubernatorial approval, presidential vote preferences, and consumer sentiment. Except for presidential vote preferences, over-time relationships are high, but they are all below the over-time stability of either the SRC or Gallup measures of partisanship (see table 6).

[10] For both the April and the July–August surveys, we conducted a separate analysis among respondents who were asked the Michigan SRC questions early in the interviews and those who were asked the SRC questions late in the interviews. Likewise, we conducted an analysis among respondents who were asked the Gallup question early in the interview and those who were asked the Gallup question late in the interview. There are occasional differences between the relationships among those asked these questions early and late in the interviews, but these differences are not systematic.

Table 3. Product-Moment Correlations between Short-Term Evaluations and Partisanship, by Question Format

	Question Format		
	Michigan SRC[a]		
Survey Dates (1992)	SRC_1	SRC_2	Gallup[b]
March 8–12:			
Presidential approval	.46 (283)	.45 (422)	.56 (264)
Gubernatorial approval	.42 (282)	.39 (421)	.40 (260)
Presidential vote preference[c]	.62 (243)	.58 (352)	.74 (233)
April 12–24:			
Presidential approval	.56 (478)	.49 (736)	.59 (464)
Gubernatorial approval	.45 (479)	.43 (743)	.49 (461)
Presidential vote preference[d]	.77 (410)	.67 (620)	.79 (402)
Consumer sentiment	.29 (430)	.28 (680)	.30 (415)
May 14–22:			
Presidential approval	.39 (256)	.35 (381)	.55 (202)
Gubernatorial approval	.40 (245)	.32 (372)	.47 (203)
Presidential vote preference[e]	.74 (163)	.70 (213)	.84 (148)
Presidential vote preference[f]	.62 (232)	.54 (351)	.75 (190)
Consumer sentiment	.29 (222)	.26 (342)	.34 (183)
July 19–August 4:			
Presidential approval	.57 (470)	.54 (699)	.61 (474)
Presidential approval (of handling economy)	.39 (483)	.38 (715)	.41 (485)
Presidential approval (of handling foreign affairs)	.49 (468)	.41 (688)	.52 (470)
Gubernatorial approval	.45 (459)	.40 (672)	.47 (457)
Presidential vote preference[d]	.77 (452)	.74 (640)	.80 (454)
Presidential vote preference[g]	.71 (467)	.68 (671)	73 (466)
Consumer sentiment	.26 (334)	.27 (500)	.31 (338)
September 12–19:			
Presidential approval	.56 (221)	.55 (334)	.61 (209)
Presidential approval (of handling economy)	.33 (221)	.35 (333)	.49 (207)
Presidential approval (of handling foreign affairs)	.37 (214)	.35 (328)	.45 (206)
Gubernatorial approval	.52 (213)	.42 (322)	.57 (201)
Presidential vote preference[d]	.83 (195)	.76 (297)	.79 (196)
Presidential vote preference[h]	.75 (209)	.69 (318)	.70 (197)
Consumer sentiment	.31 (140)	.37 (219)	.48 (145)
October 20–26:			
Presidential approval	.64 (228)	.59 (326)	.78 (207)
Presidential approval (of handling economy)	.50 (225)	.47 (324)	.55 (211)

Table 3 (*Continued*)

	Question Format		
	Michigan SRC[a]		
Survey Dates (1992)	SRC$_1$	SRC$_2$	Gallup[b]
Presidential approval (of handling foreign affairs)	.51 (219)	.46 (319)	.53 (206)
Gubernatorial approval	.48 (221)	.48 (316)	.50 (203)
Presidential vote preference[e]	.85 (177)	.83 (248)	.93 (181)
Presidential vote preference[f]	.78 (210)	.75 (307)	.87 (204)
Consumer sentiment	.41 (183)	.40 (267)	.42 (179)

Source.—See table 1.

Note.—Relationships are Pearson's r's. All correlations have been calculated so that the predicted zero-order relationships are positive. The number of cases on which these correlations are based is in parentheses.

[a] SRC$_1$ is based on respondents who identify as either Republicans or Democrats. SRC$_2$ also includes Independents who feel closer to the Republican or the Democratic party.

[b] The Gallup measure is based on respondents who say they are either Republicans or Democrats.

[c] Bush versus "any Democrat."

[d] Bush versus Clinton.

[e] Choice among Bush, Perot, and Clinton, with Perot supporters excluded.

[f] Choice among Bush, Perot, and Clinton, with Perot supporters in an intermediate position between Bush and Clinton supporters.

[g] Choice among Bush, Perot, and Clinton, if Perot were on the ballot. Perot supporters in an intermediate position between Bush and Clinton supporters.

[h] Choice among Bush, Perot, and Clinton, if Perot was still a candidate. Perot supporters in an intermediate position between Bush and Clinton supporters.

paired Bush against Clinton. In subsequent surveys, we paired Bush against Clinton, but the questions were altered to take into account whether Perot was running. For our final four surveys, therefore, we show relationships between partisanship and a two-way choice between Bush and Clinton and a three-way measure that includes support for Perot.

In five of our six surveys, Gallup partisanship is more strongly related to presidential vote choices than either of the two SRC measures. However, our September results do not support our thesis. Regardless of whether we used a two-way or three-way measure of presidential vote preference, Gallup partisanship is less related to presidential vote preferences than is SRC$_1$.

In all five surveys employing the consumer sentiment questions, Gallup partisanship is more strongly related to consumer sentiment than is either SRC measure. However, the differences are substantial only in the September survey.

These results are remarkably consistent. In 30 of the 33 comparisons, responses to the Gallup measure are more strongly related to the short-term indicators than are responses to SRC_1. In all 33 comparisons, the responses to the Gallup measure are more strongly related to these short-term indicators than are responses to SRC_2. These results strongly suggest that the Gallup question is more likely to capture short-term preferences than the Michigan SRC question is.[11]

Despite the consistency of these findings, the Gallup and SRC results are not dramatically different in statistical terms. Of course, all of the zero-order relationships reported in table 3 are statistically significant, regardless of which measure is used. To determine whether the differences in the relationships between the Gallup measure and short-term indicators are significantly different from the relationships between Michigan SRC partisanship and these indicators, we regressed these short-term indicators against a composite partisanship measure, a dummy variable for question type, and an interactive term. Assessing the statistical significance of the interactive term provides a direct test of the differences in question type (see Hanushek and Jackson 1977). We restricted our tests to the differences in the relationships between SRC_1 and Gallup. Differences are statistically significant in only three cases.[12] In spite of the absence of statistical significance between question types, the consistency of the results strongly suggests that these two question formats do tap different properties.

Differences within the Same Surveys

As we noted above, the April survey revealed a similar aggregate-level distribution of partisanship using both the Gallup and the Michigan SRC questions. But individuals may be answering these questions differently. The first part of table 4 presents the joint distribution of responses to the Michigan SRC and the Gallup partisanship questions and classifies SRC Independent leaners as Independents. Bear in mind that all respondents were asked both the full SRC battery and the basic

---

[11] Despite the consistency of our findings, we must note that Bishop, Tuchfarber, and Smith (1992, 1993) found little difference in the relationships between respondents who are asked the Gallup question and those who are asked the Michigan SRC question.

[12] The correlation between Gallup partisanship and presidential approval is significantly greater than the correlation between SRC_1 and presidential approval in both the May and the October surveys. The correlation between Gallup partisanship and approval of Bush's handling of the economy is significantly greater than the correlation between SRC_1 and economic approval in the September survey.

Table 4. Joint Distribution of Michigan SRC Partisans and Gallup Partisans (%)

Gallup Partisanship	Michigan SRC Partisanship									
	All Independents as Independents[a]					Independent Leaners as Partisans[b]				
	Republican	Independent	Democrat	Other Response	Don't Know	Republican	Independent	Democrat	Other Response	Don't Know
Republican	25.1	1.3	.3	.3	.1	26.4	.1	.3	.3	.1
Independent	2.5	37.8	2.6	.6	.8	15.9	10.9	16.1	.6	.8
Democrat	.4	1.7	21.8	.0	.2	.4	.3	23.2	.0	.2
Other response	.1	.1	.2	.3	.0	.1	.1	.2	.3	.0
Don't know	.1	.9	.3	1.5	.8	.3	.4	.6	1.5	.8

NOTE.—Survey conducted April 12–24, 1992. All respondents were asked both the Michigan SRC questions and the Gallup question.

[a] Total percent = 99.8; N = 963; percentage same on both measures = 85.9.

[b] Total percent = 99.9; N = 963; percentage same on both measures = 61.6.

Gallup question (with the order of questions randomly distributed).[13] These questions were separated by about 12 minutes, with most of the intervening questions related to public questions about Michigan. Of course, answering these questions might lead some individuals to change their partisan preferences. Also, we cannot determine whether there would have been change if respondents were asked the same question at the beginning and end of the interview.[14] Lastly, responding to the first partisanship question might influence responses to the second. However, we think it is unlikely that either "assimilation" or "contrast" effects (see Schwarz and Sudman 1992) influence our substantive conclusions (see nn. 6, 10, and 14 above and n. 15 below).

Most respondents answer the Michigan SRC and Gallup questions the same way.[15] Reading from left to right down the diagonal of the first part of the table, we can see that 85.9 percent of the respondents are classified the same way with both questions. But there is some movement away from both parties when the Gallup question is employed, and some movement toward Independence. Eleven percent of the SRC Republicans are not Gallup Republicans, mainly because they are Independents "as of today"; only 7.6 percent of the Gallup Republicans are not SRC Republicans; 13.6 percent of the SRC Democrats are not Gallup Democrats, while only 9.5 percent of the Gallup Democrats are not SRC Democrats; 9.7 percent of the SRC Independents are not Gallup Independents, while 14.8 percent of the Gallup Independents are not SRC Independents. In the second part of table 4, we reclassify Michigan SRC Independents according to the party toward which they lean. Now there are substantial differences between the two distributions, and only 61.6 percent of the respondents are classified the same on both questions.

---

[13] To the best of our knowledge, the only other study in which respondents were asked both the Michigan SRC and the Gallup question in the same survey was a National Opinion Research Center Amalgam survey conducted in 1973 (see Green and Schickler 1993).

[14] We were concerned that in a brief interview respondents would object if they were asked the identical question twice, and we were even concerned that respondents might object to being asked both partisanship questions. For both our April and July–August surveys, we instructed the interviewers to note whether respondents commented about a similar question being asked earlier in the interview. In fact, few respondents commented on having been asked a similar question earlier in the interview. Most of the interviewers were students at Michigan State University. They were not told why two questions about partisanship were being asked. For the four remaining surveys we did not inform the interviewers that two different versions of the partisanship question were being employed.

[15] We cannot definitely dismiss the possibility that asking the first partisanship question contributes to this similarity. However, it is equally plausible that asking the first question contributes to differences between the two responses. For both the April and July–August survey we examined the joint distribution of Michigan SRC partisanship and Gallup controlling for question order. In both surveys, the percentage who were classified the same on both measures was somewhat lower when the SRC questions were asked early in the interview and the Gallup question at the end. This raises the possibility of a "contrast" effect when the SRC questions precede the Gallup question. But, as noted (n. 10 above) there is no systematic tendency for the responses to the Gallup item to be more related to short-term evaluations when the Gallup item is asked after the SRC questions.

The July–August poll provides an additional opportunity to compare responses to the SRC and Gallup questions within the same survey. The first part of table 5 presents the joint distribution of responses to the Michigan SRC and the Gallup partisanship questions and classifies SRC Independents as Independents. The questions were separated by about 15 minutes. As with the April test, most of the items separating the partisanship questions were about public policy in Michigan.

Again, most respondents answer the Michigan SRC and Gallup questions the same way. Reading from left to right down the diagonal of the first part of the table, we can see that 87.2 percent of the respondents are classified the same way with both questions. There is now movement away from the Republicans when the Gallup question is employed: 11.8 percent of the SRC Republicans are not Gallup Republicans; among SRC Democrats, only 7.7 percent are not also Gallup Democrats. In the second part of table 5, we reclassify Michigan SRC Independents according to the party toward which they lean. Now there are substantial differences between the two distributions, and only 67.3 percent of the respondents are classified the same on both questions.

In both the April and the July–August polls, individual differences result mainly from movement into and out of the Independent category. In our April survey, regardless of which measure of partisanship is employed, only four SRC Republicans say they are Democrats "as of today" and only three SRC Democrats are scored as Gallup Republicans. When we classify SRC Independent leaners as Independents in our July–August poll, only two SRC Republicans say they are Democrats "as of today" and only one SRC Democrat is scored as a Gallup Republican. If SRC leaners are classified as partisans, four SRC Republicans are scored as Gallup Democrats, and two SRC Democrats are scored as Gallup Republicans, but as both tables 4 and 5 reveal, nearly one in 10 respondents is classified as a partisan on one measure and as a self-professed Independent on the other. These differences appear to result partly from short-term political conditions, with short-term forces appearing to have a greater impact upon responses to the Gallup question.

Table 5. Joint Distribution of Michigan SRC Partisans and Gallup Partisans (%)

Gallup Partisanship	Michigan SRC Partisanship									
	All Independents as Independents[a]					Independent Leaners as Partisans[b]				
	Republican	Independent	Democrat	Other Response	Don't Know	Republican	Independent	Democrat	Other Response	Don't Know
Republican	24.4	2.0	.1	.0	.0	26.2	.1	.2	.0	.0
Independent	2.4	34.6	1.4	.1	.0	13.0	10.2	15.1	.1	.0
Democrat	.2	2.9	27.9	.2	.0	.5	.0	30.6	.2	.0
Other response	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Don't know	.6	1.7	.8	.3	.2	1.1	.7	1.3	.3	.2

NOTE.—Survey conducted July 19–August 4, 1992. All respondents were asked both the Michigan SRC questions and the Gallup question.

[a] Total percent = 99.8; N = 859; percentage same on both measures = 87.2.

[b] Total percent = 99.8; N = 859; percentage same on both measures = 67.3.

Individual-Level Stability

If the Michigan SRC measure taps long-term partisan loyalties compared with the Gallup question, responses to the SRC question should be more stable over time than responses to the Gallup question. Furthermore, differences in overtime stability should become sharper as the election approaches. By comparing responses in our four panels, we can determine the stability of responses. As table 6 shows, we divided our subsets according to the questions that were asked. The number of subsets is limited because respondents were asked the same partisanship questions in March and September and the same questions in May and October. Thus, there are never more than four possible question pairings for any over-time comparison, and for the first- and third-wave comparison and for the second- and fourth-wave comparison, there are only two possible combinations. But for all panel comparisons we can compare respondents who are always asked the SRC questions with those who are always asked the Gallup question.

We conducted analyses for alternative measures of over-time stability, using a three-by-three measure including Independents as a middle category. For all of our two-wave comparisons, we report two measures of over-time stability for each subgroup: the percentage who are classified the same way on both waves and a Pearson's *r*. For our three-wave comparisons, we report the percentage who are classified the same way in all three interviews, and for our four-wave comparison we report the percentage who are classified the same way in all four interviews.

We turn first to the two-wave comparisons. There are six possible panel comparisons, and in all six cases over-time stability is highest when respondents are asked the SRC question in both waves. This greater stability is revealed both through the correlational analysis and by examining the percentage who are classified the same way in both surveys.

There are four possible question distributions for each of the four three-wave comparisons. In all four comparisons, the percentage who are consistent across three waves is highest when respondents are asked the SRC question in all three surveys.

For the four-wave comparison, there are also four possible question distributions. Once again, the percentage who are consistent across all four waves is highest for respondents who are asked the SRC question in all four surveys. Nearly three out of four are consistent across all four interviews. Among respondents who are asked the Gallup question in all four surveys, just over three out of five are consistent across all four surveys.

As expected, the differences between the SRC and Gallup formats

Table 6. Over-Time Stability of Partisanship across Four Waves, by Question Format

	Over-Time Relationship		
	Percentage Same across Waves	Pearson's *r*	*N*
Two-wave comparisons:			
Wave 1, wave 2:			
SRC, SRC	79.5	.81	205
SRC, Gallup	77.4	.73	186
Gallup, SRC	75.9	.74	199
Gallup, Gallup	76.3	.72	190
Wave 1, wave 3:			
SRC, SRC	79.0	.79	367
Gallup, Gallup	75.2	.72	359
Wave 1, wave 4:			
SRC, SRC	81.5	.83	178
SRC, Gallup	76.2	.76	172
Gallup, SRC	72.0	.74	168
Gallup, Gallup	71.4	.67	175
Wave 2, wave 3:			
SRC, SRC	84.9	.85	172
SRC, Gallup	82.6	.83	172
Gallup, SRC	80.0	.78	160
Gallup, Gallup	82.2	.79	157
Wave 2, wave 4:			
SRC, SRC	84.5	.86	323
Gallup, Gallup	79.5	.79	302
Wave 3, wave 4:			
SRC, SRC	90.0	.89	161
SRC, Gallup	82.0	.79	161
Gallup, SRC	76.9	.80	160
Gallup, Gallup	80.1	.79	151
Three-wave comparisons:			
Wave 1, wave 2, wave 3:			
SRC, SRC, SRC	74.4	. . .	164
SRC, Gallup, SRC	70.3	. . .	155
Gallup, SRC, Gallup	66.7	. . .	165
Gallup, Gallup, Gallup	69.9	. . .	153
Wave 1, wave 2, wave 4:			
SRC, SRC, SRC	76.7	. . .	159
SRC, Gallup, Gallup	68.3	. . .	145
Gallup, SRC, SRC	64.7	. . .	150
Gallup, Gallup, Gallup	64.7	. . .	150

Table 6 (*Continued*)

	Over-Time Relationship		
	Percentage Same across Waves	Pearson's *r*	*N*
Wave 1, wave 3, wave 4:			
SRC, SRC, SRC	76.5	. . .	153
SRC, SRC, Gallup	69.7	. . .	155
Gallup, Gallup, SRC	62.5	. . .	152
Gallup, Gallup, Gallup	65.1	. . .	149
Wave 2, wave 3, wave 4:			
SRC, SRC, SRC	83.0	. . .	147
SRC, Gallup, SRC	70.5	. . .	146
Gallup, SRC, Gallup	70.5	. . .	139
Gallup, Gallup, Gallup	70.7	. . .	133
Four-wave comparison:			
Wave 1, wave 2, wave 3, wave 4:			
SRC, SRC, SRC, SRC	73.6	. . .	140
SRC, Gallup, SRC, Gallup	63.4	. . .	134
Gallup, SRC, Gallup, SRC	60.0	. . .	140
Gallup, Gallup, Gallup, Gallup	61.1	. . .	131

SOURCE.—See table 1.
NOTE.—Based on a three-way comparison of Republicans, Democrats, and Independents, with Independents in the middle position. Independent leaners are classified as Independents.

are greatest in the final wave of our study, conducted only 2 weeks before the election at the point of maximum political volatility. In six of the seven comparisons including the fourth wave, respondents who are always asked the SRC question over time are at least 10 percentage points more likely to give consistent responses than are respondents who are always asked the Gallup question.

As with our other analyses, we conducted tests of significance. We restricted our comparisons to differences in the relationships between respondents who were always asked the Michigan SRC question and those who were always asked the Gallup question. We determined whether the percentage classified as stable over time was significantly greater with the SRC measure by conducting chi-square tests with 1 *df*. Just over half of the differences between respondents asked the SRC questions across time and those who were asked the Gallup question are statistically significant at the conventional .05 level, including the difference between the four-wave SRC result and the four-wave Gallup result. In six of the seven comparisons including the fourth wave, differences between respondents asked the Michigan SRC question over time and those who are asked the Gallup question are statistically significant at the .05 level. More important, results using the consistent SRC format are more stable across all 11 sets of panel comparisons.

These results clearly support our contention that the SRC measure is more stable than the Gallup measure. Moreover, the differences become strongest in analyses including the fourth wave of our panel, conducted during the peak of the 1992 general election campaign.

Conclusions

Our findings suggest that differences in question wording (in particular, the different prefaces to the Gallup and SRC partisanship questions) may lead to different conclusions about the relative stability of partisanship. Even though the differences between the Gallup and SRC results are relatively small, four principal findings support this conclusion. First, there were fewer Republicans and more Independents with the Gallup question than with the SRC question. This suggests that the Gallup item is more responsive to the volatility of the political context. Second, responses to the Gallup measure were more related to short-term evaluations and electoral preferences than were responses to the SRC measure. This suggests that the Gallup question is more likely to respond to short-term forces than is the SRC question. Third, when both questions are used in the same survey the differences that occur are found mainly in the Independent category, with the Gallup measure appearing to respond more to short-term political conditions. Finally, our panel analyses demonstrate that responses to the Gallup question are less stable over time than are responses to the SRC question. These differences are most marked when the fourth wave of our panel is employed, suggesting that differences between the measures will be amplified or attenuated according to the volatility of the political context. During many periods the political context is tranquil, and the two-question formats may yield similar readings. However, when the political context is unstable, the Gallup question may respond more to political conditions than the SRC question responds.

Taken together, these individual-level results help to explain our finding that Gallup macropartisanship varies more over time than do aggregate levels of partisanship employing the standard SRC measure. As MacKuen, Erikson, and Stimson (1992b, p. 485) remind us, "a small number of individual-level changes (due to what are sometimes called 'period effects') can generate impressive-looking net movement on the macro scale of partisanship.'' Hence aggregate-level findings based on the Gallup question may have limited bearing on the research literature based on the Michigan SRC party identification question. However, our findings do not support Converse and Pierce's conclusion (1986) that a measure that lacks the "generally speaking" wording is merely a variant of vote intention. Even though the Gallup item has a short-term focus, the question asks about political parties, not political candidates.

Although the Gallup measure is volatile, relatively few respondents switch from the Republican to the Democratic party or move from the Democrats to the G.O.P. Most of the change results from respondents moving into and out of the Independent category. In addition, the overall proportion of Gallup Independents appears to be affected by short-term political forces. This may have serious implications for the MacKuen, Erikson, and Stimson measure since it excludes Independents. Not only does excluding the middle category increase variation, but the proportion of Independents has grown substantially during the postwar years. In Gallup surveys conducted between the end of World War II and 1964, only about one respondent in five was an Independent; since then, about three out of 10 have been Independents. As our question-wording experiments demonstrate, the Gallup measure is more volatile than the Michigan SRC measure. As the proportion of Independents increases, a Gallup measure that excludes them tends to become even more volatile. Thus, as Green and Palmquist (1992, pp. 12–13) correctly observe, "That the MacKuen et al. . . . graphs of macropartisanship seem to fluctuate enormously over time reflects the fact that their measure is not mean partisanship per se, but rather the proportion of all party *identifiers* who affiliate with the Democrats."

Students of aggregate-level change would be well advised to evaluate the robustness of Gallup macropartisanship in the face of alternative scoring procedures. Alternative measures of Gallup macropartisanship might have more over-time comparability than a percentage measure based only on partisans. For example, a mean score measure would closely track the percentage measure, but it would register substantially less variation. Employing alternative measures of Gallup macropartisanship might lead to increased confidence in the utility of the Gallup measure. But however useful Gallup macropartisanship may prove to be, scholars should remain cautious in generalizing from findings based on the Gallup measure to studies of partisanship that rely on the Michigan SRC party identification measure.

Appendix
Sampling Procedures

Our results are based on statewide surveys of voting-age Michigan residents. Households were selected using random-digit dialing procedures. The original sample frame includes all possible telephone numbers in Michigan. Blocks of phone numbers that contain no residential numbers were excluded. A random sample of numbers within blocks of working residential numbers was then selected for calling. Not all of these numbers are currently in service and some are assigned to business and other nonhouseholds. When repeated calls to a selected number resulted in all nonanswers, the number was discarded as a nonworking number. A minimum of six call attempts was made before a number was discarded as nonproductive.

Within households, a single respondent was selected among the adults (the person with the most recent birthday). If the person selected was not available at the time of the initial contact with the household, interviewers scheduled a time to call back that would be convenient to the designated respondent.

The first survey was designed as the first wave of a panel. Among households reached, the completion rate was 50.6 percent. Half the noncompletions resulted from refusals, half from normal technical problems (such as time constraints) that prevented contacting the designated respondent. While maintaining the respondent's anonymity, we determined his or her first name to allow us to recontact the same respondent. Among the 1,066 respondents interviewed, 1,023 said they would be willing to be reinterviewed.

The second survey was designed as a cross-sectional survey of Michigan residents. The completion rate was 54.6 percent. Among persons not interviewed, 56 percent resulted from refusals, and 44 percent resulted from failure to contact the designated respondent.

The third survey was the second wave of the panel survey. We attempted to contact all of the 1,023 respondents who agreed to be reinterviewed, using an unlimited number of callbacks during this 9-day period. Only 31 of these persons refused to be interviewed, but another 156 could not be contacted during the interview period. Among the 836 respondents who were interviewed, 810 agreed to be reinterviewed.

The fourth survey was designed as a cross-sectional survey of Michigan residents. The completion rate was 45.9 percent. Among persons not interviewed, 46 percent resulted from refusals, and 54 percent resulted from failure to contact the designated respondent.

The fifth survey was the third wave of our panel. We attempted to reinterview all respondents, except for the 69 who said they did not want to be reinterviewed during the first or second waves. We used an unlimited number of callbacks during this 11-day period. Among the 997 respondents we attempted to interview in the third wave, only 41 refused to be interviewed, but another 184 could not be contacted for technical reasons. Among the 1,066 respondents interviewed in March, 772 were reinterviewed. Among these 772 respondents, 743 agreed to be reinterviewed.

The final survey was the fourth wave of our panel. We attempted to reinterview all respondents originally interviewed in March, with the exception of the 98 who said they did not want to be reinterviewed during the first three waves. Among the 968 we attempted to interview for our final wave, only 32 refused to be interviewed, but another 199 could not be contacted for technical reasons.

In both the September and October panels we attempted to reinterview all respondents who were originally interviewed in March (except for those who said they did not wish to be reinterviewed), regardless of whether they were interviewed in intervening panels. Survival rates will vary somewhat depending on which panels are compared. Six hundred and one respondents were interviewed in all four waves.

The panel study was based on a stratified sample of all Michigan households with telephones. Michigan counties were grouped into three strata, according to region. As the Northern counties were oversampled, a weighting factor is needed to obtain a representative sample of Michigan. In addition, the sample can be weighted to adjust for variations in the sampling frame based on gender and race. As the different question formats randomly varied across these weighting factors, we present the results based upon the real *N*s. The cross-sectional surveys did not employ stratification by region, but weighting factors can be used to adjust the sampling frames based on gender and race. Because the two question orders randomly varied across these weighting factors, we present the actual results.

References

Abramson, Paul R. 1983. *Political Attitudes in America: Formation and Change*. San Francisco: W. H. Freeman.

Abramson, Paul R., John H. Aldrich, and David W. Rohde. 1991. *Change and Continuity in the 1988 Elections*, rev. ed. Washington, DC: Congressional Quarterly Press.

Abramson, Paul R., and Charles W. Ostrom, Jr. 1991. "Macropartisanship: An Empirical Reassessment. *American Political Science Review* 85:181–92.

———. 1992. "Question Wording and Macropartisanship." *American Political Science Review* 86:481–86.

Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69:1218–31.

Allsop, Dee, and Herbert F. Weisberg. 1988. "Measuring Change in Party Identification in an Election Campaign." *American Journal of Political Science* 32:996–1017.

Alwin, Duane F., Ronald L. Cohen, and Theodore M. Newcomb. 1991. *Political Attitudes over the Life Span: The Bennington Women after Fifty Years*. Madison: University of Wisconsin Press.

Alwin, Duane F., and Jon A. Krosnick. 1991. "Aging, Cohorts, and the Stability of Sociopolitical Orientations over the Life Span." *American Journal of Sociology* 97:169–95.

Asher, Herbert B. 1974. "Some Consequences of Measurement Error in Survey Data." *American Journal of Political Science* 18:469–85.

Bishop, George F., Alfred J. Tuchfarber, and Andrew E. Smith. 1992. "Question Wording and Partisanship: An Experimental Test of the Hypothesis." Paper presented at the annual meeting of the American Political Science Association, Chicago, September 3–6.

———. 1993. "Question Form and Context Effects in the Measurement of Partisanship: Conflicting Evidence and Rival Hypotheses." Paper presented at the annual meeting of the American Political Science Association, Washington, DC, September 2–5.

Borrelli, Stephen, Brad Lockerbie, and Richard G. Niemi. 1987. "Why the Democrat-Republican Partisanship Gap Varies from Poll to Poll." *Public Opinion Quarterly* 51:115–19.

Brody, Richard A. 1991. "Stability and Change in Party Identification: Presidential to Off-Years." In *Reasoning and Choice: Explorations in Political Psychology*, ed. Paul M. Sniderman, Richard A. Brody, and Philip E. Tetlock, pp. 179–205. Cambridge: Cambridge University Press.

Brody, Richard A., and Lawrence S. Rothenberg. 1988. "The Instability of Partisanship: An Analysis of the 1980 Presidential Election." *British Journal of Political Science* 18:445–65.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Converse, Philip E. 1976. *The Dynamics of Party Support: Cohort-Analyzing Party Identification*. Beverly Hills, CA: Sage.

Converse, Philip E., and Roy Pierce. 1986. *Political Representation in France*. Cambridge, MA: Harvard University Press.

Curtin, Richard T. 1982. "Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers." *Public Opinion Quarterly* 46:340–52.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Franklin, Charles H. 1992. "Measurement and the Dynamics of Party Identification." *Political Behavior* 14:297–309.

Green, Donald Philip, and Bradley Palmquist. 1990. "Of Artifacts and Partisan Instability." *American Journal of Political Science* 34:872–902.

———. 1992. "How Stable Is Party Identification?" Revised version of paper presented at the Roundtable "Party Identification: A New Look at the New Look?" at the annual meeting of the American Political Science Association, Chicago, September 3–6.

Green, Donald Philip, and Eric Schickler. 1993. "Multiple-Measure Assessment of Party Identification." *Public Opinion Quarterly*, vol. 57 (in press).

Hanushek, Eric A., and John E. Jackson. 1977. *Statistical Methods for Social Scientists*. New York: Academic Press.

Hunter, John E., and T. Daniel Coggin. 1976. "A Reanalysis of Achen's Critique of the Converse Model of Mass Political Belief." *American Political Science Review* 70:1226–29.

Kagay, Michael R. 1991. "Rise in GOP Loyalty Now Ebbing, Poll Finds." *New York Times*, June 2.

Kohut, Andrew. 1991. "Questioning Party Identification." *Public Perspective* 2 (September–October): 21–22.

Lockerbie, Brad. 1989. "Change in Party Identification: The Role of Prospective Economic Evaluations." *American Politics Quarterly* 17:291–311.

MacKuen, Michael B., Robert S. Erikson, and James A. Stimson. 1989. "Macropartisanship." *American Political Science Review* 83:1125–42.

———. 1992a. "Moving Attractors and the Character of Partisan Change." Paper presented at the annual Political Methodology. Conference, Cambridge, MA, July 16–19.

———. 1992b. "Question Wording and Macropartisanship." *American Political Science Review* 86:475–83, 485–86.

Miller, Warren E. 1992. "Rethinking the Logic of Causal Inference." Paper presented at the Roundtable "Party Identification: A New Look at the New Look?" at the annual meeting of the American Political Science Association, Chicago, September 3–6.

Niemi, Richard G., and M. Kent Jennings. 1991. "Issues and Inheritance in the

Formation of Party Identification." *American Journal of Political Science* 35:970–88.

Ostrom, Charles W., Jr., and Dennis M. Simon. 1988. "The President's Public." *American Journal of Political Science* 32:1096–1119.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.

Schwarz, Norbert, and Seymour Sudman, eds. 1992. *Context Effects in Social and Psychological Research*. New York: Springer-Verlag.

Stimson, James A. 1991. *Public Opinion in America: Moods, Cycles, and Swings*. Boulder, CO: Westview.

Weisberg, Herbert F., and Charles E. Smith, Jr. 1991. "The Influence of the Economy on Party Identification in the Reagan Years." *Journal of Politics* 53:1077–92.

DOES SIMILARITY BREED RESPECT? INTERVIEWER EVALUATIONS OF MEXICAN-DESCENT RESPONDENTS IN A BILINGUAL SURVEY

AÍDA HURTADO

Abstract A guiding assumption of survey methodology is that similarity between interviewers and respondents on important social characteristics increases the validity of the information obtained in the interview. Therefore, an effort is made in surveys to match interviewers and respondents on important social characteristics such as race. This study examines the interviewer evaluations of Spanish/English bilingual Mexican descendants from a national survey. The results show that the respondents' social characteristics were related to interviewer evaluations even though interviewers were of Mexican descent, Spanish/English bilinguals, and knowledgeable of Mexican and U.S. culture. Respondents who met conventional societal criteria of success (being better educated, having higher incomes, being fair skinned and European looking, and speaking English) were evaluated as better able to understand questions, better able to answer questions, and more cooperative and interested in the interview. However, this correlational study could not determine whether these evaluations were based on differences in respondents' abilities, on bias in favor of higher-status respondents, or on some combination of the two. Further research needs to determine the causal direction of these interviewer evaluations and their possible implications for the training of interviewers, especially in bilingual surveys with predominantly working-class populations who are racially distinct.

People engaged in social interaction constantly have to interpret a person's character on the basis of salient social characteristics, such as gender, race, and social class. The face-to-face survey interview is no exception as interviewers' and respondents' evaluations of each other inevitably enter the situation (Hyman et al. 1954; Singer and

AÍDA HURTADO is associate professor of psychology at the University of California, Santa Cruz.

Public Opinion Quarterly Volume 58:77–95 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5801-0006$02.50

Kohnke-Aguirre 1979). Precisely because social evaluations cannot be avoided, many studies of survey methodology have examined the impact of interviewer characteristics on respondents (Finkel, Guterbock, and Borg 1991; Hutchinson and Wegge 1991; Lueptow, Moser, and Pendleton 1990; Weeks and Moore 1981). Race of interviewer has been investigated most often. Other visible interviewer characteristics, such as gender and age, have been studied less, although initial research shows they are somewhat influential since "whenever respondents classify interviewers into a larger category and believe they know what would offend or please persons in that category," the interview process is affected (Schuman and Kalton 1985, p. 686).

The impact of respondents' characteristics on interviewers has received much less attention, although almost 40 years ago Hyman et al. (1954) noted the problem of interviewer role expectations in survey research:

> We might conceive of role expectations to denote the tendencies of interviewers to believe that certain attitudes or behaviors occur in individuals of given group memberships, and therefore expect answers of a certain sort from particular persons. Some of these beliefs might well occur because of traditional role prescriptions characteristic of all societies, as illustrated in G's remarks: "I just don't think the average woman has as much social consciousness as the average man." Some role expectations might well be posited on the basis of an oversimplified belief, a stereotype about some ethnic group. In either case, at the initial moment of interaction in the interview, the respondent might be pigeon-holed on the basis of some membership cue, and the structure of his attitudes would be expected to correspond with that role. [Pp. 61–62]

The theory of status characteristics and expectation states (Berger, Rosenholtz, and Zelditch 1980) suggests the characteristics that are easily activated in performance situations. These *diffuse status characteristics*, such as gender, age, race or ethnicity, and social class, have two critical features: *(a)* status of the characteristic (black vs. white, Latino vs. Anglo, male vs. female, middle class vs. working class) is differentially valued in society; and *(b)* each status arouses an expectation of competence or incompetence at socially valued tasks (Cohen 1981). Diffuse status characteristics are easily aroused, especially when *(a)* the performance of the person being evaluated affects the success of the evaluator, who *(b)* is motivated to succeed and *(c)* believes it legitimate and necessary to take the other person's qualities and behavior into account. These conditions apply to the survey interview situation, and interviewers are likely to expect some respondents to be more competent than others (Weiss 1968–69).

The study reported here investigates the relationships between four respondent characteristics on the interviewer's evaluations of respondents' ability to understand the questions, respondents' ability to express themselves, and respondents' cooperation and interest in the interview. The data are provided by interviewers' evaluations of respondents in a national survey of the Mexican-descent population in the United States.

However, although the present investigation started from the hypothesis of interviewer bias, I cannot exclude the possibility that the causal sequence is partly or entirely different: that the ability to understand questions and to express oneself, as well as cooperation and interest, occur because of real variations in ethnicity and the other background factors I have measured. The present study is basically correlational, and only indirectly can I point to evidence that interviewer bias is a major causal factor. I return to this issue in the conclusions.

Respondent Characteristics

Two respondent characteristics that fit the criteria for diffuse status characteristics are *gender* and *socioeconomic status*. Sociolinguistic research has demonstrated that judges differentially evaluate the verbal fluency and intellectual competence of women and men and of middle-class and working-class (or poor) subjects. Men's speech is judged to be more assertive and fluent than that of women (Thorne and Henley 1975); working-class people's speech is judged less articulate than that of upper-middle-class people (Giles and Farrar 1979; Giles and Sasson 1983). In these studies, the content of the messages is held constant and only the speech style is varied between experimental conditions. However, there is the possibility in actual interviews that better-educated individuals with higher incomes may indeed make better respondents because they are used to expressing their opinions and are more familiar with the interview situation.

Two other characteristics, *whether the respondents speak English or Spanish* and *how Anglicized their facial features are*, also arouse expectations of competence and have special significance for surveys of Latino populations. Sociolinguistic research has documented that evaluation of a person's level of motivation and ability are affected by what language they speak. In general, in the United States, English speakers are judged more positively than people who speak Spanish, and people who speak standard English as more intelligent and attractive than people who speak English with a Spanish accent (Ryan and Carranza 1977). Again, in these studies the content of the messages remains the same and researchers vary only the languages used by the speakers. In the interview situation, however, the respondents who speak English may be more familiar with the concepts and etiquette appropriate for face-to-face interviews and therefore may be evaluated more positively than Spanish speakers because they are, in fact, better able to understand questions and articulate answers.

Language alone is not what distinguishes the Mexican-descent population in this country. This population is a product of the racial commingling of native Mexican Indians, blacks, and various European groups. The mixture has produced a range of phenotypes—fair-skinned, European-looking individuals as well as dark-skinned, Indian-looking people (Arce, Murguia, and Frisbie 1987; Telles and Murguia 1990, 1992). Phenotype affects socioeconomic attainment and other life chances such that blacks who are light skinned and Chicanos who are fair and European looking have higher socioeconomic status than darker blacks and darker Chicanos who are Indian looking (Arce, Murguia, and Frisbie 1987; Hughes and Hertel 1990; Mullins and Sites 1984; Ransford 1970; Relethford et al. 1983; Telles and Murguia 1990, 1992). In the United States, the phenotype of the respondent is likely to be related to an interviewer's role expectations (Pettigrew 1964, p. 7) in ways similar to other diffuse status characteristics. This study examined two aspects of the respondent's phenotype—skin color and facial features.

The potential relationship between language and phenotype is especially important for national surveys in the United States. Even bilingual, ethnically matched interviewers may hold different performance expectations for respondents who speak a language other than English and do not look European. These expectations may then affect the course of the interview, causing interviewers to probe more thoroughly with respondents judged to be more competent and motivated or to wait longer for them to respond. To date, bilingual surveys have not been used to examine how language and other characteristics of non-English speakers affect the reactions of survey interviewers (Grimshaw 1969).

Spanish speakers present an especially serious problem for national surveys. Spanish speakers are already a significant language minority in the United States. In 1980, over 11 million people, representing about 5 percent of the United States population, reported speaking Spanish in the home. No language other than English can claim as many speakers as Spanish (U.S. Bureau of the Census, Department of Commerce 1983). Moreover, the issue of how to deal with Spanish speakers in national surveys will become more important as increased immigration from Latin American countries and population growth among those of Mexican descent produce a larger proportion of Spanish speakers in the United States (Hurtado et al. 1992). Representativeness is jeopardized by the current survey practice of eliminating Spanish or other non-English speakers, usually because the survey team does not include bi- or multilingual interviewers. Declaring English the "official language" does not solve the problem (Hurtado 1990). Representativeness simply cannot be achieved in many areas of the country when non-English speakers are deleted from the sample.

This research note begins the process of establishing how language and other respondent characteristics may be related to interviewer's evaluations in surveys covering the Latino population, including those who are monolingual Spanish speakers. It provides some preliminary evidence of the possible impact of respondent characteristics on interviewers who were themselves all of Mexican descent, bilingual, and carefully trained and supervised.

Research Questions

The following questions are addressed: *(a)* How were the respondents' socioeconomic status, gender, and phenotype related to the interviewers' evaluations of the respondents' ability to understand the questions and to articulate answers and their cooperation and interest in the interview? *(b)* Were these evaluations related to whether the interview was conducted in Spanish or English? *(c)* Did these evaluations change over the course of the interview? and *(d)* Did the relationship of language (or phenotype) depend on social class?

Method

SAMPLE

The data came from the 1979 survey of persons of Mexican descent conducted by the Institute for Social Research of the University of Michigan. The survey was based on a probability sample of Mexican-ancestry households in the southwestern United States (California, New Mexico, Arizona, Texas, and Colorado) and in the Chicago metropolitan area. In its final composition, the sample is representative of almost 90 percent of the total U.S. population of Mexican ancestry identified in the 1970 U.S. census.

Eligible respondents were heads of households and/or their spouses who were at least of half Mexican ancestry. When only one adult was eligible (i.e., of Mexican origin), that person was interviewed. When both the household head and spouse were eligible, one was randomly selected. Nearly 11,000 households were screened for ancestry in 41 primary sampling units; of these, almost 1,300 were of Mexican descent and 991 interviews were completed. Further details of the population definition and multistage sampling design are found in Santos (1985). Other than the U.S. census, this survey represents the most comprehensive data source on people of Mexican descent. The interview format itself covered labor force issues, language, culture, social identity, political consciousness, mental health, and family life.

The sample is relatively young, on average 40 years of age.[1] Average income is $11,446, and on average the respondents did not go beyond junior high school. Two-thirds of the respondents are female, and the majority were currently married (75 percent), the rest either separated, divorced, or widowed (19 percent), and only a few never married (6 percent). The majority of respondents were born in the United States (54 percent of the men and 67 percent of the women). Compared to Bureau of the Census estimates for the Mexican-descent population (derived from the Current Population Survey [CPS] for March 1979), Mexican descendants in the sampled households were, on average, 1 year younger and had approximately one-half year less schooling. Households in the sample were slightly larger; they had an average of 4.4 persons compared to 4.1 in the CPS. Except for gender, differences between the sample and the census profile were, for the most part, negligible (Arce, Murguia, and Frisbie 1987).

INTERVIEWER CHARACTERISTICS AND INTERVIEW CONTEXT

A substantial amount of research guides the attempts of survey researchers to increase interviewer-respondent rapport (Schuman and Kalton 1985). Thus, this study selected interviewers who were Spanish/English bilinguals and had knowledge of Mexican as well as U.S. culture. All had at least 1 year of interviewing experience, and a 3-day training session was used to teach procedures of screening respondents for eligibility and conducting this particular interview. The interview booklets were printed in English and Spanish, side by side. English and Spanish equivalency was obtained by independent translators back-translating from English to Spanish and vice versa. Disagreements in translation were settled by the bilingual research team.

MEASURES: INDEPENDENT VARIABLES

Respondent choice of Spanish or English for the interview. To help determine the respondent's language preference, the interviewers were asked to pay close attention to language cues as the respondent opened the door (e.g., to which radio and television stations residents of the household were listening or what language children were speaking). The interviewer was asked to follow the language cues coming from the household environment to guide which language was used in the initial greeting. If no language cues were present, the interviewer used English and switched to Spanish if the respondent looked perplexed. If the respondent's preference was still uncertain, the interviewer asked directly, "What language do you feel most at ease using, English or Español?" A total of 518 (52 percent) of the respondents took the interview in Spanish and 473 (48 percent) in English.

---

[1] The standard deviation for average age is 14.81 years and for average income $7,593. Average years of schooling is 8.39 with a standard deviation of 4.47.

Respondent socioeconomic status (SES). The respondent's years of education and family income in the preceding year were combined to construct SES. For income, the respondent was asked which of 16 categories, ranging from less than $2,000 to more than $30,000, fit the family income. For education, respondents were asked how many years of schooling they had received. The responses ranged from none to postgraduate education. The SES index was composed from two separate scales. The first scale coded number of years of education from 1 to 6 (1 = <6 years, 2 = 6 years, 3 = 7–9 years, 4 = 10–11 years, 5 = high school, 6 = more than high school). Similarly, the second scale coded the respondents' income into six categories (1 = <$6,000, 2 = $6,000–9,000, 3 = $10,000–14,999, 4 = $15,000–19,999, 5 = $20,000–24,999, 6 = >$25,000). The SES index combined the two scales by adding them and dividing by 2. (The distribution for the SES variable is given in the Appendix.)

Phenotype. The interviewers evaluated the respondent's *skin color* and *facial features* on two separate scales ranging from 1 to 5, with a high score on both scales indicating fair complexion and European appearance and a low score indicating dark complexion and Indian-looking facial features.

Respondent gender was noted by the interviewer.

MEASURES: DEPENDENT VARIABLES

These measures, which are summarized in table 1, come from the now-standard procedure in face-to-face interviews of having the interviewer complete an evaluation of the respondent and the interaction shortly after leaving the respondent's home. The data are typically used to ascertain why there is inconsistency in responses or why the respondent may have given incoherent answers to open-ended questions. Aside from their practical use in facilitating more accurate coding, such data can be used to answer substantive questions as raised here. One of the main reasons these data are often ignored for research purposes is that interviewers may not record their observations immediately after the interview, thus dramatically decreasing these data's

Table 1. Distribution of Dependent Variables

	Percent
Interviewer evaluation of the respondent's verbal ability:	
Please describe the respondent's ability to express himself/herself on the scale below:	
1. Limited vocabulary; expresses himself/herself with great difficulty	10.2
2.	29.3
3.	37.8
4. Excellent vocabulary; very articulate	18.5
9. Missing	4.2
Interviewer evaluations of the respondent's ability to understand questions:	
The respondent's understanding of the question was:	
1. Poor	7.5
2. Fair	22.9
3. Good	40.3
4. Excellent	23.8
9. Missing	5.5
Interviewer evaluation of the respondent's motivation:	
The respondent's attitude at the beginning of the interview was:	
1. Uncertain, nervous	20.0
2. Neutral, relaxed	28.9
3. Cooperative, helpful	47.8
9. Missing data	3.3
The respondent's attitude at the end of the interview was:	
1. Less cooperative, helpful	4.8
2. No change from the beginning of the interview	53.8
3. More cooperative, helpful	38.1
9. Missing data	3.3
The respondent's interest at the beginning of the interview was:	
1. Not very interested	10.4
2. Somewhat interested	55.8
3. Very interested	30.5
9. Missing data	3.3
The respondent's interest at the end of the interview was:	
1. Less interested	6.5
2. No change from the beginning of the interview	54.1
3. More interested	35.9
9. Missing data	3.5

NOTE.—$N = 991$.

reliability. However, in the National Chicano Survey there was a team leader in the field for every three interviewers whose sole responsibility was to monitor on a daily basis the quality of the survey materials, including the completion of the interviewer observations. Therefore, unlike other major surveys in which interviewers complete these observations just before they send the materials to the home office and days after the actual interview, on-site team leaders in this study insured that this was done at the end of each day.

Interviewer evaluation of the respondent's verbal ability. The interviewers evaluated respondents' ability to articulate answers to the survey questions (verbal fluency). The scale for the verbal fluency judgment was reversed so that a high score indicates excellent verbal ability.

Interviewer evaluation of the respondent's ability to understand questions. Interviewers evaluated the respondent's apparent comprehension of the survey's questions on a scale ranging from 1 to 4, which was reversed so that a high score indicates excellent understanding.

Interviewer evaluation of the respondent's motivation. The respondent's cooperation and interest in the interview at the beginning and end of the interview were evaluated on two scales, each ranging from 1 to 3. On each of these scales, the interviewer evaluated how motivated the respondent was at the beginning of the interview and whether the respondent became more or less motivated or stayed the same by the end of the interview. The scales were reversed so that a high score on the first scale indicates that the respondent initially was very interested and had a very cooperative attitude. A high score on the second scale indicates that the respondent had become more cooperative and more interested over the course of the interview.

DATA ANALYSIS

Multiple classification analysis (MCA), a multivariate technique directly comparable to regression that examines relationships between categorical independent variables and a single dependent variable (Andrews et al. 1973), was used to determine which of the independent variables were related to the interviewer evaluations after adjusting for the effects of the other independent variables.

Results

The overall results are summarized in table 2. Table 2 presents the MCA betas. Since MCA betas do not indicate direction, table 2 also

Table 2. MCA Analysis of Interviewer Evaluations of Respondent's Language Skills and Motivation

	Respondent's Understanding of the Questions[a]		Respondent's Ability to Express Herself/Himself[a]		Respondent's Attitude at Start of the Interview[b]		Respondent's Interest at Start of the Interview[c]	
	Betas	Means	Betas	Means	Betas	Means	Betas	Means
Language of interview	.10*		.02		.06		.08*	
1. English		1.94		1.69		2.23		2.16
2. Spanish		1.76		1.66		2.32		2.26
Respondent's gender	.02		.00		.06		.03	
1. Male		1.83		1.67		2.33		2.19
2. Female		1.86		1.67		2.24		2.22
Respondent's SES	.45*		.46*		.22*		.28*	
1. Low		1.18		1.03		2.09		1.99
2.		1.51		1.31		2.00		2.05
3.		1.67		1.48		2.23		2.10
4.		1.96		1.78		2.34		2.31
5.		2.28		2.08		2.47		2.40
6. High		2.53		2.46		2.56		2.51

Respondent's skin color	.12*		.17*		.07		.05	
1. Very dark		1.78		1.57		2.26		2.27
2.		1.72		1.51		2.21		2.20
3.		1.87		1.67		2.34		2.22
4.		1.88		1.73		2.23		2.20
5. Very light		2.06		2.02		2.36		2.22
Respondent's facial features	.16*		.21*		.11*		.05	
1. Indian looking		1.67		1.48		2.31		2.30
2.		1.70		1.44		2.15		2.19
3.		1.87		1.71		2.34		2.21
4.		1.93		1.81		2.24		2.20
5. European looking		2.26		2.18		2.50		2.30
Multiple *R* (adjusted)	.54		.52		.22		.23	
Variation explained (R^2 [%])	29		29		5		6	

NOTE.—MCA betas are significant (indicated with an asterisk) if they are at least twice as large as the standard error from a multiple regression using the same predictors and dependent measures.

[a] 1 = poor; 4 = excellent.

[b] 1 = uncertain, nervous; 3 = cooperative, helpful.

[c] 1 = not interested; 3 = interested.

presents the adjusted means for various categories of the predictor variables in order to show how each predictor relates to the dependent variables after adjusting for the effects of the other predictors. Table 3 summarizes analyses of change in the interviewers' evaluations of the respondents' motivation. It depicts their evaluations of whether the respondents became more or less motivated or stayed the same *after* adjusting for the evaluations they made of the same respondents at the beginning of the interview.

The first question is, How were the four respondent characteristics—SES, gender, phenotype, and language of interview—related to interviewer evaluations? Overall, SES yielded the strongest relationship between these characteristics, and gender had the weakest relationship.[2] A second major question concerned evaluations of respondent motivation at the beginning and end of the interview and whether the relationships between respondent characteristics waned over the course of the interview. In general, the relationship of socioeconomic status persisted, while the relationship of language of interview and phenotype were less pronounced at the end.

Turning to specific results, the respondents' SES was significantly related to all evaluations the interviewer was asked to make. Interviewers rated higher-status respondents as better able to understand the survey questions than lower-status respondents. Higher-status respondents also were judged as better able to express themselves. Because these evaluations of language skills were made *after* a long verbal interaction with the respondent, the SES relationship might merely reflect the objectively greater verbal competency of respondents with more education and higher family incomes. Analyses of the motivation evaluations suggest, however, that at least some bias could have been involved in the interviewer evaluations of lower- and higher-status respondents. As noted above, the interviewers were asked how cooperative and interested the respondents had been at the *beginning* of the interview. The fact that the higher-SES respondents were rated as more motivated at the onset of the interview suggests that the interviewers' initial expectancies could have been somewhat class biased. (Substantiation of an initial class bias would be stronger, of course, had the interviewers been asked to make this evaluation at the start of the interview.) The results also show that the interviewers rated higher-SES respondents as having *improved* more than other respon-

---

[2] Unfortunately, we did not measure the interviewers' SES and phenotype. All interviewers were Spanish/English bilinguals with no apparent preference for one language or another. However, we did conduct MCAs using the interviewers' gender as one additional predictor and there were no significant results.

Table 3. MCA Analysis of Interviewer's Evaluations of Respondent's Cooperation and Interest at the End of the Interview Controlling for Initial Rating

	Cooperation at the End of Interview[a]		Interest at the End of Interview[b]	
Interviewer Evaluations	MCA Betas[c]	Means	MCA Betas[d]	Means
Cooperation at start of interview	.29*			
1. Uncertain, nervous		2.28		
2.		2.43		
3. Cooperative, helpful		2.91		
Interest at start of interview			.15*	
1. Not very interested				2.22
2.				2.31
3. Very interested				2.53
Language of interview	.05		.07	
1. English		2.37		2.35
2. Spanish		2.32		2.27
Respondent's gender	.05		.02	
1. Male		2.31		2.29
2. Female		2.37		2.32
Respondent's SES:				
1. Low	.12*	2.34	.10*	2.30
2.		2.33		2.29
3.		2.29		2.31
4.		2.32		2.25
5.		2.39		2.36
6. High		2.48		2.37
Respondent's skin color	.06		.12*	
1. Very dark		2.28		2.36
2.		2.35		2.27
3.		2.32		2.27
4.		2.38		2.28
5. Very light		2.40		2.48
Respondent's facial features	.09		.10*	
1. Indian looking		2.31		2.23
2.		2.26		2.23
3.		2.37		2.35
4.		2.39		2.31
5. European looking		2.35		2.40

NOTE.—MCA betas are significant (indicated with an asterisk) if they are at least twice as large as the standard error from a multiple regression using the same predictors and dependent measures.

[a] 1 = less cooperative; 4 = more cooperative.

[b] 1 = less interested; 3 = more interested.

[c] Multiple R (adjusted) = .33; variation explained (R^2) = 13 percent.

[d] Multiple R (adjusted) = .22; variation explained (R^2) = 7 percent.

dents. The interviewers were asked whether the respondents' level of cooperativeness and interest had lessened, stayed the same, or increased over the course of the interview. The means presented in table 3, which are adjusted for the relationships of other respondent characteristics and for initial interviewer evaluations of motivation, indicate that respondents on average were evaluated more cooperative and interested at the end, and that this was significantly more true of higher-status respondents.

Phenotype is also related to the interviewer evaluations but the two components of phenotype—skin color and facial features—have different relationships to the dependent variables. Whereas the respondents' facial features are related to the interviewer evaluations of comprehension of questions, articulation of answers, and cooperation, skin color is only related to comprehension of questions and articulation of answers. These relationships persist even when SES is held constant. Holding the relationship of SES constant with the other predictors is important because among Mexican-descent persons, as among blacks, there is a positive correlation between SES, fair coloring, and European-looking features (Arce, Murguia, and Frisbie 1987; Hughes and Hertel 1990; Mullins and Sites 1984; Ransford 1970; Telles and Murguia 1990). The means in table 2 suggest a somewhat linear relationship between phenotype and evaluations of comprehension of questions and articulation of answers: the lighter and more European looking the respondents, the more they were judged as understanding the survey's questions and able to express themselves. In contrast, the phenotype relationship on evaluations of cooperation was nonlinear and held only for facial features: it was only the most extreme category, comprising the 12 percent who were considered the most "European-looking," that were evaluated as significantly more cooperative than other respondents in the interview. The interviewers drew no consistent motivational distinctions among respondents who were distributed across the rest of the two phenotype scales and were not, for example, unusually critical of the motivation of respondents who were "very dark" or the most "Indian-looking."

Table 3 shows that this subtle, positive regard for the most Anglicized respondents was not evident in the interviewers' views of final cooperation, although fair and European-looking respondents were judged as having become more *interested* than others over the course of the interview.

The relationship of language of interview was less obvious. Spanish-interviewed respondents were evaluated as more motivated at the start of the interview, although English-interviewed respondents were judged better able to understand the interview questions. Moreover, the more positive judgment of those who elected an English interview persisted even after adjusting for their higher SES.[3]

The respondents' gender, in contrast, had a very small relationship to the interviewers' evaluations. The mean ratings indicate that women were considered less cooperative than men at the beginning of the interview although the MCA beta was not significant (.06).

A last question was whether the evaluations of either phenotype or language were related to the respondents' social class. Phenotype and language can be considered markers of ethnicity, and, as such, they may interact with social class. Dark-skinned, Indian-looking people are likely to draw stereotyped evaluations in this society, particularly when they also have working-class standing. I tested the possible interactions: between SES and skin color, SES and facial features, and SES and language of interview. This was done by conducting two sets of MCAs. One followed the standard assumptions of the MCA program, that the relationships under examination can be understood in terms of an additive model and "that the average score (on the dependent variable) for a set of individuals is predictable by *adding together* the 'effects' of several predictors" (Andrews et al. 1973, p. 18). The other was carried out using a pattern variable, which is a combination of the categories of two predictors (SES and skin color, e.g.). Comparing the adjusted *R* from the analysis with a pattern variable with the adjusted *R* from the additive analysis gives a test for interaction. If the differences between the two *R*'s is statistically significant, then there is evidence of an interaction. The results were clear-cut. None of the possible interactions proved statistically reliable. These ethnic markers and SES were related to the interviewers' evaluations, but in an additive fashion.

Discussion

The social characteristics of the respondents were related to the interviewers' evaluations even when other significant variables were held constant. In general, there is a significant relationship between respondents who meet conventional societal criteria of success (being better educated, having higher incomes, being fair skinned and European looking, and speaking English) and interviewers' positive evaluations of their being better able to understand the interview questions, being more articulate, and being more cooperative and interested in the interview. It is unclear from these correlational results, however, whether all of the respondent characteristics indeed result in biased evaluations or whether these respondents are more competent respondents.

---

[3] One interpretation is that differences in questionnaire translation affected linguistic equivalency, especially in difficult topics like mental health. There were no significant differences in how interviewers rated the difficulty of the topics in the questionnaire according to whether respondents took the interview in English or in Spanish. Regardless of language of interview, three areas (political consciousness, family-related matters, and mental health) were mentioned as the most problematic for respondents.

The interviewers who conducted this survey were carefully selected, culturally sensitive, bilingual, and well trained. Therefore, it is not surprising that the relationships were moderate. The strongest relationship involved SES and, to a lesser extent, respondent phenotype. The strong relationship between SES and the interviewer evaluations may, in fact, reflect the respondents' competencies as a result of education and income. Another possibility may lie in the characteristics of the interviewers themselves, all of whom had at least a high school education and some of whom had a college degree. The similarity in education between the higher-SES respondents and the interviewers may have influenced their assessments.

The results for skin color and facial features are a little more difficult to explain as being the result of objective differences in the respondents' abilities. There is no reason for fair-skinned, European-looking respondents to be better able to understand the questions and to articulate answers when the variation in their SES is held constant. Again, similarity between these respondents and the interviewers may have influenced the interviewers' evaluations. Having participated in all of the interviewer training sessions, I also would consider most of the interviewers as light skinned and generally having European facial features. The interviewers' level of education and income is also consistent with Telles and Murguia's (1990, 1992) findings that better-educated Mexican descendants from higher economic backgrounds are more Anglo looking in skin color and facial features. Therefore, certain differences between interviewers and certain categories of respondents were inadvertently "built in" to the research design.

It cannot be determined from these data whether the interviewer evaluations were actually biased. However, the fact that higher-SES and European-looking respondents were considered more motivated at the onset of the interview suggests some bias may have been involved. These effects diminished at the end of the interview, especially for those social characteristics emphasized during training (such as gender and language). The differential evaluations made by the interviewers persisted for characteristics that were not directly addressed in the training, such as SES and phenotype, and that, in the case of SES, may reflect real differences in ability between respondents. As a precaution against the possibility of bias, greater attention could be given to SES and phenotype in the training of interviewers. However, such training would have to handle subtle issues. As mentioned earlier, it was comprehension of questions, *not* verbal fluency, that was judged more highly for the English speakers. Most of the respondents grew up in predominantly Spanish-speaking homes (83 percent of the 991 respondents stated that *Spanish* or *mostly Spanish* was spoken in the home when they were growing up) and may have Spanish-accented English. However, their English skill may be related to the interviewers' evaluations of comprehension even when their English fluency is hampered. A potential solution for this is to train interviewers to recognize the relationship of language to the *intellectual* expectations of respondents. The phenotype evaluations were also subtle. There was not obvious discrimination against dark-skinned, Indian-looking respondents but a positive evaluation of extremely fair, European-looking respondents. Again, training may help interviewers to recognize this finely differentiated judgment.

This analysis is primarily descriptive. Controlled experimentation on the effects of respondent-interviewer characteristics identified here is needed before definitive conclusions may be drawn. At the very minimum, however, these correlational results may provide a clue that other researchers may follow to confirm that matching interviewers and respondents on ethnicity and language alone may not keep interviewers from being guided by their *role expectations* of class and phenotype. One reason that the percentage of variance explained by this analysis is small (between 29 percent and 6 percent) may be precisely because interviewers were carefully selected, trained, and supervised. Thus, if bias is at the root of these differential judgments, these relationships are likely to be even stronger when interviewers and respondents are *not* matched for ethnicity and language and when the survey addresses issues of race and ethnicity (Hatchett and Schuman 1975–76; Reese et al. 1986; Weeks and Moore 1981; Welch, Comer, and Steinman 1973). Training interviewers about *role expectations* may be especially crucial when conducting bilingual surveys with predominately working-class respondents who are racially distinct. Future research should include phenotypic ratings and sociodemographic characteristics of interviewers to fully decipher the pattern of correlational results obtained in this study.

Appendix

Table A1. Distribution for the SES Variable

Code Values	Percent
1–1.5	21.7
2–2.5	26.4
3–3.5	20.5
4–4.5	18.0
5–5.5	9.6
6	3.8
99	9.2

NOTE.—*N* = 991.

References

Andrews, Frank M., James N. Morgan, John A. Sonquist, and Laura Klem. 1973. *Multiple Classification Analysis*. Ann Arbor: Institute for Social Research, University of Michigan.

Arce, Carlos H., Edward Murguia, and W. Parker Frisbie. 1987. "Phenotype and Life Chances among Chicanos." *Hispanic Journal of Behavioral Sciences* 9:19–32.

Berger, J., S. J. Rosenholtz, and M. Zelditch, Jr. 1980. "Status Organizing Processes." *Annual Review of Sociology* 42:479–508.

Cohen, C. E. 1981. "Person Categories and Social Perception: Testing Some Boundaries of the Processing Effects of Prior Knowledge." *Journal of Personality and Social Psychology* 40:441–52.

Finkel, Steven E., Thomas M. Guterbock, and Marian J. Borg. 1991. "Race-of-Interviewer Effects in a Preelection Poll: Virginia 1989." *Public Opinion Quarterly* 55(3):313–30.

Giles, Howard, and Kathryn Farrar. 1979. "Some Behavioral Consequences of Speech and Dress Styles." *British Journal of Social and Clinical Psychology* 6(2):209–10.

Giles, Howard, and Caroline Sasson. 1983. "The Effect of Speaker's Accent, Social Class Background and Message Style on British Listener's Social Judgments." *Language and Communication* 3(3):305–13.

Grimshaw, Allen D. 1969. "Language as an Obstacle and as Data in Sociological Research." *Items* 23:17–21.

Hatchett, Shirley, and Howard Schuman. 1975–76. "White Respondents and Race-of-Interviewer Effects." *Public Opinion Quarterly* 39:523–28.

Hughes, M., and B. R. Hertel. 1990. "The Significance of Color Remains—A Study of Life Chances, Mate Selection, and Ethnic Consciousness among Black Americans." *Social Forces* 68(4):1105–20.

Hurtado, Aída. 1990. "English-Only Advocates' Attitudes toward Language Diversity in the United States: What Functions Do They Serve?" Working paper, University of California, Santa Cruz, Bilingual Research Group.

Hurtado, Aída, David Hayes-Bautista, Robert Burciaga Valdez, and Anthony C. Hernandez. 1992. "Redefining California: Latino Engagement in a Multicultural Society." Working paper, University of California, Los Angeles, Chicano Studies Research Center.

Hutchinson, Kevin L., and David G. Wegge. 1991. "The Effects of Interviewer Gender upon Response in Telephone Survey Research." *Journal of Social Science Behavior and Personality* 6(3):573–84.

Hyman, Herbert, William J. Cobb, Jacob J. Feldman, Clyde W. Hart, and Charles Herbert Stember. 1954. *Interviewing in Social Research*. Chicago: University of Chicago Press.

Lueptow, Lloyd B., Susan L. Moser, and Brian F. Pendleton. 1990. "Gender and Response Effects in Telephone Interviews about Gender Characteristics." *Sex Roles* 22(1–2):29–42.

Mullins, Elizabeth I., and Paul Sites. 1984. "The Origins of Contemporary Eminent Black Americans: A Three-Generation Analysis of Social Origin." *American Sociological Review* 49:672–85.

Pettigrew, Thomas F. 1964. *A Profile of the Negro American*. Princeton, NJ: Van Nostrand.

Ransford, H. E. 1970. "Skin Color, Life Chances, and Anti-white Attitude." *Social Problems* 18:164–78.

Reese, Stephen D., Wayne A. Danielson, Pamela J. Shoemaker, Tsan-Kuo Chang, and Huei-Ling Hsu. 1986. "Ethnicity-of-Interviewer Effects among Mexican-Americans and Anglos." *Public Opinion Quarterly* 50:563–72.

Relethford, J. H., M. P. Stern, S. P. Gaskill, and H. P. Hazuda. 1983. "Social Class, Admixture, and Skin Color Variation in Mexican-Americans and Anglo-Americans Living in San Antonio, Texas." *American Journal of Physical Anthropology* 61:97–102.

Ryan, Ellen B., and Miguel A. Carranza. 1977. "Ingroup and Outgroup Reactions to Mexican American Language Varieties." In *Language Ethnicity and Ingroup Relations*, ed. Howard Giles. London: Academic Press.

Santos, Robert. 1985. "A Methodological Report on the Sample Design of the 1979 National Chicano Survey." Working paper series, no. 11, Stanford Center for Chicano Research.

Schuman, Howard, and Graham Kalton. 1985. "Survey Methods." In *The Handbook of Social Psychology*, ed. Gardner Lindzey and Elliot Aronson, 1:635–98. New York: Erlbaum.

Singer, Eleanor, and Luane Kohnke-Aguirre. 1979. "Interviewer Expectation Effects: A Replication and Extension." *Public Opinion Quarterly* 43:245–60.

Telles, E. E., and E. Murguia. 1990. "Phenotypic Discrimination and Income Differences among Mexican Americans." *Social Science Quarterly* 71(4):682–93.

———. 1992. "The Continuing Significance of Phenotype." *Social Science Quarterly* 73(1):120–22.

Thorne, Barrie, and Nancy Henley. 1975. *Language and Sex: Difference and Dominance*. Rowley, MA: Newberry House.

U.S. Bureau of the Census, Department of Commerce. 1983. *Condition of Hispanics in America Today*. Washington, DC: Government Printing Office.

Weeks, Michael F., and R. Paul Moore. 1981. "Ethnicity-of-Interviewer Effects on Ethnic Respondents." *Public Opinion Quarterly* 45:245–49.

Weiss, Carol H. 1968–69. "Validity of Welfare Mothers' Interview Responses." *Public Opinion Quarterly* 32:622–33.

Welch, Susan, John Comer, and Michael Steinman. 1973. "Interviewing in a Mexican-American Community: An Investigation of Some Potential Sources of Response Bias." *Public Opinion Quarterly* 37:115–26.

DIMINISHED LIFETIME SUBSTANCE USE OVER TIME

AN INQUIRY INTO DIFFERENTIAL UNDERREPORTING

MICHAEL FENDRICH
CONNIE M. VAUGHN

Abstract This study investigates underreporting of lifetime marijuana and cocaine use in the National Longitudinal Survey of Youth. Comparing responses provided in 1988 with responses provided in 1984, this study evaluates the prevalence and correlates of two indicators of underreporting, including use denial and use reduction. At follow-up in 1988, a high rate of underreporting was evident for both marijuana and cocaine, with rates of use reduction observed for just under one-third of all 1984 marijuana users and just over one-third of all 1984 cocaine users. Outright denial of use in 1988 was almost twice as prevalent among cocaine users as it was among marijuana users; nearly 19 percent of all cocaine users denied use at follow-up, compared with nearly 12 percent of all marijuana users. Correlates of underreporting varied by substance and by measure of underreporting. The most consistent correlates of underreporting were interview mode, race/ethnicity, and educational status. Those interviewed by telephone, minority respondents, and those with lower levels of education were more likely to underreport. Race/ethnicity effects were particularly striking in our analyses. Black respondents had at least twice the odds of underreporting compared with white/other respondents for every indicator of underreporting. Findings are compared with other recent research on underreporting and are discussed in the context of recent substance use prevalence findings.

MICHAEL FENDRICH is assistant professor of psychology, Department of Psychiatry, Institute for Juvenile Research, University of Illinois at Chicago. CONNIE M. VAUGHN is a product development specialist at Information Resources, Inc., in Chicago. At the time this article was written, she was a research specialist in the Department of Psychiatry, Institute for Juvenile Research, University of Illinois at Chicago. This research was supported in part by a grant from the National Institute on Drug Abuse (grant no. 5R29DA07995-02) and by a grant from the University of Illinois at Chicago. The authors would like to acknowledge the programming assistance provided by Yanchun Xu.

Public Opinion Quarterly Volume 58:96–123 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5801-0002$02.50

Introduction

Denial of symptomatology upon reinterview among the previously symptomatic has been a consistent finding in psychiatric epidemiology. Researchers have concluded that a reinterview results in "retest artifact" or a decrease in the willingness of previously symptomatic individuals to disclose socially undesirable or sensitive information (Bromet et al. 1986; Henderson, Byrne, and Duncan-Jones 1981; Jorm, Duncan-Jones, and Scott 1989; Rubio-Stipec et al. 1992). Retest artifact calls into question the validity of inferences about socially undesirable information drawn from longitudinal research. Prior research suggests that the more sensitive the question being asked, the more pronounced the retest artifact will be (Jorm, Duncan-Jones, and Scott 1989). Such findings may have considerable implications for the validity of longitudinal survey research conducted on another sensitive topic—substance use.

The present analyses investigate retest artifact in responses to survey questions about lifetime use of marijuana and cocaine over two waves of interviews in the National Longitudinal Survey of Youth (NLSY), including both the 1984 and the 1988 surveys. We address the following questions: Among respondents identifying themselves initially as users in 1984, to what extent do respondents report reduced levels of lifetime use at follow-up? Among this same subsample of 1984 lifetime users, to what extent is use denied in 1988? What respondent and interview mode characteristics are associated with use reduction and denial? Do observed patterns and correlates of use reduction and denial differ for marijuana and cocaine?

Literature Review

At least four previous longitudinal analyses have evaluated reinterview effects in substance use surveys. Mensch and Kandel (1988) compared responses to questions about marijuana use over two waves (1980 and 1984) in the NLSY and found that 9 percent of those initially reporting use indicated that they never used marijuana at follow-up 4 years later. These authors found that blacks and Hispanics and those with less than a high school education were more likely to deny previously reported substance use. They also noted that "experimental" or low-level users were more likely to deny previous substance use. Fendrich et al. (1992) analyzed responses over 8 weeks to questions about quantity and frequency of drug use among self-identified drug users in an ethnographic sample. These authors found a general tendency toward reporting of reduced use over a series of eight prospective interviews. The pattern held across substances, and was consistent across ethnic/racial groups. Anglin, Hser, and Chou (1993) investigated the long-term (10-year) consistency of retrospective drug use reports covering overlapping time periods in a sample of male opiate addicts committed to a compulsory drug treatment program by the California courts. The findings of the Anglin et al. study were not suggestive of a retest artifact; compared to the initial interview, the 10-year follow-up interview yielded higher levels of retrospective use for an overlapping period than did the initial interview. Collins et al. (1985) compared retrospective reports of substance use provided over two waves of interviews by high school youths who participated in a smoking prevention program. On the second wave of interviews, there was an overall tendency for youths to recall less use of cigarettes and alcohol than they reported at the initial interview. This study suggested that recency of use may be an important variable influencing the level of retrospective use at follow-up; higher levels of retrospectively reported marijuana use were associated with more recent use at follow-up.

It should be noted that most of the studies demonstrating retest artifact in general populations have focused on psychiatric symptomatology and illness and not substance use per se (although substance abuse diagnoses are considered). Three large psychiatric studies of adult samples demonstrated patterns of symptom reduction among the previously symptomatic (Bromet et al. 1986; Eaton et al. 1989; Rice et al. 1987). One recent study of children, adolescents, and young adults (Fendrich and Warner, in press) found that patterns of lifetime symptom report reduction were limited to internalizing disorders such as depression and were not in evidence for symptoms of conduct disorder and substance abuse.

Aquilino and LoSciuto (1990) argue that since questions about drug use are "threatening" to respondents they are affected by "socially desirable" response sets that bias toward underreporting. According to Sudman and Bradburn (1974), questions that evoke socially desirable response sets are sensitive to interview mode. Accordingly, interview modes favoring more privacy (the self-report) or more subject awareness of and control over the interview process (the face-to-face interview) should be conducive to honest reporting about substance use. In contrast, an interview mode that limits a respondent's sense of privacy or control over the interview process—such as the telephone interview—should not be conducive to honest disclosure about substance use.

Several recent cross-sectional studies have shown that telephone administration of drug use questionnaires leads to considerable underreporting of drug use when compared with self-administered questionnaires or face-to-face interviews (Aquilino 1992; Aquilino and

LoSciuto 1990; Gfroerer and Hughes 1991, 1992; Johnson, Hougland, and Clayton 1989). Aquilino and LoSciuto (1990) argue that self-administered questionnaires minimize response effects due to social desirability. One recent study analyzing the National Household Survey suggested that respondents were less willing to disclose use of substances to interviewers than they were in self-administered questionnaires (Turner, Lessler, and Devore 1992). Analyzing the 1988 NLSY data, Schober et al. (1992) found lower rates of marijuana and cocaine use for private face-to-face interviews than for self-reports. They found that mode effects varied according to respondent race/ethnicity; in contrast to other respondents, black respondents reported similar prevalence rates of cocaine use across interview modes. Interestingly, they also found *lower* rates reported for marijuana and cocaine use in private face-to-face interviews than in nonprivate face-to-face interviews.

The above findings about ethnicity differences are at odds with studies comparing telephone interviews with face-to-face interviews. For example, Aquilino and LoSciuto (1990) compared substance use reporting in telephone interviews to self-administered questionnaires for blacks and whites and found significant underreporting in telephone interviews only for black respondents. Similarly, in a study comparing substance use reports given in face-to-face interviews with reports given in telephone interviews, Aquilino (1992) concluded that underreporting for the latter condition was significantly higher for black respondents. Note that this parallels the finding noted earlier by Mensch and Kandel (1988). It also parallels other research suggesting underreporting by black respondents on sensitive questions about abortion (Jones and Darroch-Forrest 1992). We found one study that investigated the interaction between gender and response mode effects; Johnson, Hougland, and Moore (1991) found that underreporting of substance use in telephone surveys was limited to male respondents. Additional findings reported by Mensch and Kandel (1988) suggest the possible importance of another interview mode variable in the study of underreporting in longitudinal drug use research: interviewer familiarity. Mensch and Kandel (1988) found that multiple contacts with the same interviewer depressed rates of reporting about cocaine.

Despite evidence that mode and subject characteristics may contribute to substance use underreporting, prior research has not simultaneously examined both types of variables. In previous cross-sectional research investigating interview mode effects, underreporting was evaluated in the aggregate for the group exposed to the specific interview condition. Rates of substance use reports were compared across groups defined by similar respondent characteristics. Since these prior studies were cross-sectional and used only a single assessment of substance use, individual-level indicators of underreporting could not be constructed. Without individual-level indicators of underreporting, respondent characteristics could not be examined as direct correlates of this behavior. In prior studies, sociodemographic variables were examined as interaction effects; investigators examined the variation in sensitivities to interview mode across demographic subgroups. Longitudinal studies assessing lifetime substance use over multiple waves facilitate the construction of individual level indicators of underreporting. In longitudinal studies, an index evaluating the difference between substance use frequency reports over two waves of interviews can be constructed for each individual respondent; since lifetime use is cumulative, diminished lifetime use over time can serve as an individual-level indicator of underreporting. Such individual-level estimates facilitate the direct examination of the association between underreporting and any subject characteristic of interest; they also facilitate the examination of the relative importance of different variables as predictors of underreporting in multivariate models.

Methods

SAMPLE

The NLSY is a nationally representative sample of 12,686 men and women who were ages 14–21 years when they were first interviewed in 1979 (Center for Human Resources Research 1991). The study was designed so as intentionally to oversample blacks and Hispanics as well as economically disadvantaged whites. A supplemental sample of young adults drawn from active military duty was also collected. Twelve waves of interviews were conducted on this sample between 1979 and 1990. Retention rates consistently approached or exceeded 90 percent in each of the survey years. Questions about illicit substance use were asked in three of the survey years—1980, 1984, and 1988. In 1988, half of the subjects were randomly assigned to drug use supplements that were administered by an interviewer and half of the subjects were randomly assigned to drug use supplements that were completed by self-report.

The focus of the current study is on the responses to surveys administered in 1984 and 1988, two surveys that included nearly identical questions about lifetime use for two illicit substances, cocaine and marijuana. Since the major emphasis of this report is on drug use recall (retest artifact) and its correlates, all analyses are based on unweighted samples. This study employs two subsamples of the NLSY for all analyses. The first subsample consists of all respondents who completed questions about marijuana use in 1984 and 1988 who also reported lifetime use of marijuana in 1984 (N = 6,204). The second subsample consists of all respondents who completed questions about cocaine use in 1984 and 1988 who also reported lifetime use of cocaine in 1984 (N = 1,589).

VARIABLES

Analyses investigated a number of sociodemographic and survey condition variables in association with retest artifact. We included a variable comparing three categories of race/ethnicity (blacks, Hispanics, and whites/other), a six-level indicator of 1988 income, an indicator of employment status, and an indicator of education (highest level of schooling completed).[1] Indicators of age and sex of respondent were also included in the analyses.

In order to address the influence of sociocultural norms on underreporting behavior we included two other sociodemographic variables, urban versus rural residence in 1988 and marital status in 1988. Prior research suggests that there have been consistently lower rates of reported substance use in rural areas as compared with urban areas (Johnston, O'Malley, and Bachman 1988). Prior research also suggests that marriage and the adoption of "conventional" lifestyles is associated with a reduction in reported levels of use (see Newcomb and Bentler 1988; Yamaguchi and Kandel 1984).

Interview condition variables included an indicator of interviewer familiarity (the number of times a subject had contact with the 1988 interviewer),[2] as well as a categorical indicator of interview condition reflecting the following four mutually exclusive possibilities: interviewed face-to-face alone, interviewed face-to-face in the presence of another person, interviewed over the telephone, and interviewed through self-administered questionnaire. We employed two indicators of drug use involvement in our analyses, including an assessment of initial levels of use in 1984 (using a four-category indicator that was identical across substances) and an indicator of the most recent time of use for each particular drug; recent users (those who used marijuana or cocaine within the past year in 1984) were contrasted with all other users (those who used marijuana or cocaine more than 1 year prior to the 1984 interview).

---

[1] The income variable included five levels of income; we also included an additional category reflecting missing responses on this question.

[2] Interviewer identification codes from the 1979 and 1980 surveys were not retrievable. This variable thus measures the number of contacts with the 1988 interviewer since 1980.

ANALYSES

Two dimensions of retest artifact were explored, including reports of reduced lifetime use at follow-up (use reduction) and reports of nonuse at follow-up (use denial). Because of the truncated and skewed distribution of the use frequency variables at each interview (see below), we treated use reduction as a categorical indicator with all respondents reporting a lifetime reduction at follow-up coded as 1 and all others coded as 0. Cross-tabulations and logistic regression models were constructed to evaluate correlates of use reduction and denial for both marijuana and cocaine. Within regression models, interaction terms were constructed in order to evaluate the consistency of mode effects within certain subgroups. Sex and age of respondent were included as control variables in all multivariate analyses.

Results

Table 1 describes the demographic characteristics of the two user subgroups included in the analyses. Each of the subgroups is compared with all other "nonusers" in 1984 for that substance. Statistically significant differences between users and nonusers of marijuana and cocaine were observed with respect to all of the demographic variables except highest level of education. Users of both substances were more concentrated in the older age groups (26–32 years old) than nonusers. A significantly lower percentage of users compared with nonusers of cocaine and marijuana were black, Hispanic, female, married, and from rural residences. Subgroup differences were more pronounced for cocaine use than they were for marijuana use reports. Nevertheless, the distributions on each of the demographic variables for the two subgroups of lifetime users were roughly parallel, with the same rank ordering for nearly every variable category within each subgroup. Table 1 suggests that there is somewhat limited power to detect differences within each of the main demographic variables; this is especially true for race, marital status, and residence.

Tables 2 and 3 compare substance use frequency reports across the two interview periods, 1984 and 1988. The bold areas in each table indicate underreporting in 1988 (i.e., they show frequencies for respondents whose lifetime use showed a reduction at follow-up). By summing up the total number of respondents that are bold and dividing this number by the total number of 1984 users for each substance, rates of underreporting can be generated. For marijuana (table 2), 31.2 percent (1,937/6,204) of the sample underreported their 1984 use upon reinterview. For cocaine, 39 percent (618/1,589) of the sample underre-

Table 1. Selected Sociodemographic Characteristics by Drug Use Status in 1984

	1984 Marijuana User (%)		1984 Cocaine User (%)	
	No	Yes	No	Yes
1988 age (years):				
23–25	32.9	28.2	31.1	25.0
26–28	39.1	40.2	39.2	42.4
29–32	28.0	31.5	29.7	32.6
N	3,836	6,204	8,405	1,589
Race/ethnicity:				
Black	29.8	24.3	28.3	16.4
Hispanic	19.2	14.4	16.6	14.5
White/other	51.0	61.3	55.1	69.0
N	3,836	6,204	8,405	1,589
Sex:				
Female	58.5	46.4	53.1	40.0
Male	41.5	53.6	46.9	60.0
N	3,836	6,204	8,405	1,589
1988 marital status:				
Never married	38.3	39.1	37.9	43.2
Married	50.6	45.4	48.4	41.9
Other[a]	11.1	15.5	13.7	14.9
N	3,834	6,204	8,401	1,589
Residence:				
Rural	25.8	17.6	22.2	12.8
Urban	74.2	82.4	77.8	87.2
N	3,672	5,925	8,017	1,535
Highest education completed:				
< high school	17.7	17.7	17.9	16.3
High school graduate	43.7	45.3	45.0	43.4
Some college	20.6	20.8	20.4	22.0
College graduate	18.0	16.3	16.7	18.3
N	3,802	6,158	8,339	1,574

NOTE.—Since all data are unweighted, these percentages are not population estimates; these percentages relate only to respondent groups defined in the study. Differences between users and nonusers of marijuana and cocaine were significant at the .001 level for all demographic variables except highest education completed.

[a] Divorced, widowed, or separated.

Table 2. 1984 and 1988 Lifetime Marijuana Use Frequency among 1984 Marijuana Users ($N = 6{,}204$)

1988 Marijuana Use Frequency (No. of Times Ever)	1984 Marijuana Use Frequency (No. of Times Ever)[a]							
	1–9 ($N = 2{,}631$)		10–39 ($N = 1{,}107$)		40–99 ($N = 777$)		100+ ($N = 1{,}689$)	
	%	*N*	%	*N*	%	*N*	%	*N*
0	**19.3**	**508**	**7.6**	**84**	**6.0**	**47**	**5.3**	**89**
1–9	55.1	1,450	**30.4**	**337**	**15.1**	**117**	**10.7**	**181**
10–39	15.3	403	37.7	417	**24.4**	**190**	**10.4**	**176**
40–99	5.1	135	12.4	137	25.0	194	**12.3**	**208**
100+	5.1	135	11.9	132	29.5	229	61.3	1,035
Underreporting	19.3	508	38.0	421	45.6	354	38.7	654

Note.—Sample included all marijuana users from 1984 with 1988 completed interviews and with completed answers on 1984 use frequency question. Numbers in bold represent frequencies for respondents whose lifetime use showed a reduction at follow-up.

[a] These categories were recoded to create consistent categories across the interview periods.

Table 3. 1984 and 1988 Lifetime Cocaine Use Frequency among 1984 Cocaine Users (N = 1,589)

1988 Cocaine Use Frequency (No. of Times Ever)	1984 Cocaine Use Frequency (No. of Times Ever)[a]							
	1–9 (N = 792)		10–39 (N = 404)		40–99 (N = 194)		100+ (N = 199)	
	%	N	%	N	%	N	%	N
0	**23.7**	**188**	**14.9**	**60**	**10.3**	**20**	**16.6**	**33**
1–9	48.6	385	**30.6**	**124**	**19.1**	**37**	**14.6**	**29**
10–39	17.0	135	36.9	149	**28.4**	**55**	**14.6**	**29**
40–99	5.5	44	10.6	43	21.1	41	**21.6**	**43**
100+	5.1	40	6.9	28	21.1	41	32.7	65
Underreporting	23.7	188	45.5	184	57.7	112	67.3	134

Note.—Sample included all cocaine users from 1984 with 1988 completed interviews and with completed answers on 1984 use frequency question. Numbers in bold represent frequencies for respondents whose lifetime use showed a reduction at follow-up.

[a] These categories were recoded to create consistent categories across the interview periods.

ported their 1984 use upon reinterview. Total denial of use was considerably higher for cocaine (18.9 percent) than for marijuana (11.7 percent). Although those in the lowest use (or experimental use) category were the modal group of deniers for both marijuana and cocaine, our findings are at odds with the suggestions of Mensch and Kandel (1988): underreporting is prevalent across use categories, with increasingly higher rates of underreporting observed for those at higher initial use levels; underreporting is not a phenomenon limited to experimental users. Indeed, for cocaine, the highest rate of underreporting is observed among the group of the heaviest 1984 users. Over two-thirds of those reporting 100 or more times of use in 1984 reported less use at follow-up. Even more striking, this same group showed the second highest rate of total denial of any use group (16.6 percent). Thus, denial appears to be a more serious problem for cocaine use than it is for marijuana use. Underreporting of use levels is consistently observed across substances and increases with increasing levels of initial use.

The differences between reports at each wave of interviews are underscored by statistical comparisons with respect to mean category levels among marijuana and cocaine users. Overall mean differences with respect to levels of use indicated slight but highly significant drops in reported marijuana and cocaine use between interviews (mean difference for marijuana, .21; mean difference for cocaine, .29; both paired t-tests are significant at the .001 level). Effect size estimates for the overall difference suggest only a very slight or small effect for both reductions (see Cohen 1988; effect sizes are both less than .3). An inspection of the correlation between use levels reported at each wave suggests that there is more consistency in reported use levels for marijuana than there is for cocaine. The correlation between use level reports for marijuana was .59, compared with .37 for cocaine.

Table 4 displays the results of bivariate comparisons across groups according to reductions in use level. All variables except for age and sex showed some level of association with use reduction or denial for at least one substance.

Race/ethnicity comparisons. Race/ethnicity showed consistent effects across outcome measures and across substances. Blacks showed the highest rate of use reduction and denial, followed by Hispanic respondents, and then by respondents in the white/other group. The largest differences for race were observed in the contrast between black respondents and white/other respondents on the denial outcomes: 33.2 percent of the black respondents denied cocaine use at follow-up compared with 15 percent of the white respondents.

Residence and marital status comparisons. Residence was associated with underreporting only for cocaine and not for marijuana. Residents of rural areas were more likely to report less use and more likely to deny use of cocaine than residents of urban areas (17.7 percent of the urban respondents denied cocaine use at follow-up compared with 25 percent of the rural respondents). Expectations about the role of marital status in underreporting were only confirmed for cocaine use. Married respondents denied cocaine use at a higher rate than those who were never married and than those who were divorced, widowed, or separated (rates of denial were 22.1 percent, 17.5 percent, and 14.3 percent, respectively, for those classified as married, never married, and other). Although the absolute magnitude of the difference in rates was relatively small, a significantly elevated percentage of use reduction for marijuana was reported by those who were divorced, widowed, or separated (rates of use reduction were 34.6, 30.9, and 30.3 percent, respectively, for those classified as other, never married, and married).

Employment, income, and education comparisons. Employment status was significantly associated with marijuana use denial; those who were out of the labor force or unemployed reported slightly elevated rates of use denial for marijuana compared with those who were employed or in the military (rates of denial were 14.6, 11.2, 12.3, and 10.5 percent, respectively, for those who were out of the labor force, employed, unemployed, and in the military). With respect to cocaine,

Table 4. 1988 Use Reduction and Denial among 1984 Users

	Substance Reported On					
	Marijuana			Cocaine		
	N	% Reporting Less Use	% Denying Use	*N*	% Reporting Less Use	% Denying Use
Interview mode:						
Self-report	2,964	31.3	10.4	742	36.5	16.7
Interview (alone)	2,606	31.0	12.5	657	39.1	18.6
Interview (someone present)	262	26.7	10.3	77	41.6	22.1
Interview (phone)	309	35.9	20.1	94	54.3	36.2
		n.s.	***		**	***
Interviewer familiarity (1988):						
Single contact	4,286	30.7	11.3	1,150	37.4	17.2
> 1 contact	1,904	32.4	12.8	435	42.8	23.7
		n.s.	n.s.		n.s.	**
Race/ethnicity:						
Black	1,508	40.9	19.3	261	48.3	33.2
Hispanic	891	36.5	14.9	231	39.4	20.8
White/other	3,805	26.1	8.0	1,097	36.6	15.0
		***	***		***	***
1988 age (years):						
23–25	1,752	29.9	11.7	397	38.8	21.4
26–27	1,667	33.4	12.4	451	40.8	19.7
28–29	1,567	30.2	11.4	404	35.9	17.6
30–32	1,218	31.6	11.3	337	40.1	16.6
		n.s.	n.s.		n.s.	n.s.

Table 4 (*Continued*)

	Substance Reported On					
	Marijuana			Cocaine		
	N	% Reporting Less Use	% Denying Use	*N*	% Reporting Less Use	% Denying Use
Sex:						
Male	3,327	31.1	11.3	954	40.4	19.4
Female	2,877	31.4	12.2	635	36.7	18.3
		n.s.	n.s.		n.s.	n.s.
1988 income:						
≤ 12,000	1,244	37.3	15.0	305	43.3	19.7
12,001–20,000	1,030	31.6	11.1	241	41.9	20.3
20,001–30,000	1,033	29.3	10.6	261	33.3	16.5
30,001–43,000	914	25.5	10.2	245	42.9	22.4
43,001 +	903	27.5	9.0	270	33.3	12.2
Missing	1,080	33.7	13.3	267	38.6	22.6
		***	***		*	*
1988 highest education:						
< high school	1,088	38.7	15.4	256	49.2	26.6
High school graduate	2,790	31.8	11.6	683	40.4	18.9
Some college	1,278	29.3	11.3	347	37.5	18.7
College graduate	1,002	24.0	8.3	288	28.1	12.8
		***	***		***	***

1988 employment:						
Employed	4,708	29.8	11.2	1,221	38.4	18.3
Unemployed	391	34.0	12.3	101	40.6	22.8
Out of the labor force	952	36.7	14.6	233	38.6	19.7
Military	153	32.7	10.5	34	52.9	26.5
		**	*		n.s.	n.s.
1988 marital status:						
Never married	2,422	30.9	11.7	687	35.8	17.5
Married	2,818	30.3	11.8	665	41.8	22.1
Other[a]	962	34.6	11.6	237	39.7	14.3
		*	n.s.		n.s.	*
Residence:						
Urban	4,882	31.4	11.7	1,339	37.6	17.9
Rural	1,043	29.5	11.7	196	44.9	25.0
		n.s.	n.s.		*	*
Most recent 1984 use:						
Use in last year	2,090	36.1	7.8	1,002	40.5	17.3
No use in last year	4,101	28.7	13.7	579	36.1	21.8
		***	***		n.s.	*
1984 lifetime use frequency (no. of times):						
1–9	2,631	19.3	19.3	792	23.7	23.7
10–39	1,107	38.0	7.6	404	45.5	14.9
40–99	777	45.6	6.0	194	57.7	10.3
100+	1,689	38.7	5.3	199	67.3	16.6
		***	***		***	***

---

[a] Divorced, widowed, or separated.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

nonsignificant trends suggested elevated rates of use reduction and denial for those in the military and those who were unemployed compared with those who were employed or out of the labor force. Income was associated with use reduction and denial for both marijuana and cocaine. For marijuana, those earning under $12,000 and those missing on the income question showed elevated rates of reported use reduction and denial compared with those in other income groups. The relationship between income and underreporting was more complicated for cocaine use reporting. Only two income groups had cocaine use reduction rates considerably under 40 percent—those in the middle income group and those in the highest income group. These same two groups also had the lowest rates of denial (with rates for other groups approaching or exceeding 20 percent). In contrast with income, the associations between education and underreporting for both marijuana and cocaine use were clearly monotonic. The higher the education level, the lower the rate of use reduction and denial for both marijuana and cocaine.

Interviewer familiarity and interview mode. Multiple contacts with the 1988 interviewer showed a significant association with underreporting only for cocaine use; 23.7 percent of the respondents with multiple contacts with the 1988 interviewer denied cocaine use compared with 17.2 percent of the respondents with one contact. Interview mode was significantly associated with denial of marijuana and cocaine use; it was associated with reports of use reduction for cocaine. Denial of marijuana and cocaine use was slightly lower for those who completed self-reports than for those who completed face-to-face interviews. The largest and most consistent differences in interview mode effects were observed for those who were interviewed over the telephone. Respondents interviewed over the telephone had twice the rate of denial for marijuana and cocaine compared with repondents who completed self-reports. Over half of the respondents interviewed over the telephone reported less cocaine use at follow-up (54.3 percent); over one-third of the respondents interviewed over the telephone denied cocaine use at follow-up (36.2 percent).

Use recency comparisons. The relationship between use recency in 1984 and underreporting varied according to outcome measure. Compared with other respondents, those reporting marijuana use during the past year in 1984 (recent users) were significantly more likely to report reduced lifetime use in 1988 (36.1 percent of the recent users reported reduced use, compared with 28.7 percent of all other respondents). On the other hand, those reporting recent use were significantly less likely to deny use of marijuana or cocaine in 1988. Those not reporting recent marijuana use in 1984 had nearly twice the rate of denial of recent marijuana users (rates of denial were 7.8 percent for those with recent marijuana use, compared with 13.7 percent for those without recent use). Those not reporting recent cocaine use during 1984 had slightly higher rates of denial than recent cocaine users (rates of denial were 21.8 percent for those without recent cocaine use, compared with 17.3 percent for those with recent use). It should be pointed out that use recency in 1984 is significantly associated with use frequency for both marijuana and cocaine. Those who used more recently in 1984 were also more likely to report a higher level of lifetime use in 1984. The contrasting effects of recency on denial and use reduction parallel those found for the use frequency variable.

Tables 5 and 6 evaluate the impact of the covariates discussed above simultaneously in a set of logistic regressions. For each covariate, odds ratios provide an estimate of the effect of category membership (vs. nonmembership) on the relative odds of each outcome measure (see Hosmer and Lemeshow 1989). Findings are summarized for each substance below.[3]

Marijuana use reduction and denial. Significant effects for use reduction and denial in marijuana were detected for indicators of interview mode, race, income, educational status, marital status, and 1984 use frequency (see table 5). Compared with those who were interviewed face-to-face without anyone else present as the reference group, those who were interviewed by telephone were significantly more likely to deny and report reduced use of marijuana. Compared with white/other respondents, black and Hispanic respondents were significantly more likely to deny and report reduced use of marijuana. Compared with those with less than a high school level of education, those with higher levels of education were significantly less likely to deny and report reduced use of marijuana. Similarly, those in the higher income groups were significantly less likely to deny and report reduced use than those earning $12,000 or less. Compared with those who were married, those who were never married were significantly less likely to deny and report reduced use of marijuana. Two additional indicators were significant predictors of marijuana use reduction: gender and use recency. Women were more likely to report reduced use of marijuana than were men. Those who reported marijuana use in 1984 were less likely to report reduced use compared with other respondents. Finally, note that the signs of the use frequency variables were reversed for use reduction and denial; frequency of use in 1984

---

[3] Dummy variables for employment status were omitted from regression models for two reasons. Respondents with missing values on any variable were eliminated from regression analyses; these resulted in too few cases for two employment dummy variable categories (military and unemployed) to provide reliable parameter estimates. Additionally, these indicators were highly collinear with other indicators of social status that were already in the model (education and income).

Table 5. Logistic Regressions: Predicting Use Reduction and Denial of Marijuana Use (N = 5,846)

	Use Reduction			Use Denial		
Variable	b	Standard Error	Odds Ratio	b	Standard Error	Odds Ratio
1988 interview mode:						
Face-to-face/alone	. . .	. . .	1.00	. . .	. . .	1.00
Face-to-face/someone present	−.30	.16	.74	−.25	.24	.78
Self-report	.06	.06	1.06	−.18	.09	.84
Telephone	.34	.14	1.40*	.73	.17	2.08***
1988 interviewer familiarity:						
One contact	. . .	. . .	1.00	. . .	. . .	1.00
> 1 contact	.06	.06	1.06	.01	.09	1.02
Race:						
White	. . .	. . .	1.00	. . .	. . .	1.00
Black	.83	.07	2.29***	.96	.10	2.61***
Hispanic	.60	.09	1.82***	.58	.12	1.78***
Age (years):						
24–25	. . .	. . .	1.00	. . .	. . .	1.00
26–27	.09	.08	1.09	.13	.11	1.14
28–29	−.02	.08	.98	.07	.11	1.06
30–32	.04	.09	1.04	.03	.13	1.02
Sex:						
Male	. . .	. . .	1.00	. . .	. . .	1.00
Female	.21	.06	1.24***	−.04	.09	.97

Income:						
< 12,000	. . .	. . .	1.00	. . .	. . .	1.00
12,001–20,000	−.14	.10	.87	−.32	.14	.73*
20,001–30,000	−.22	.10	.80*	−.34	.15	.71*
30,001–43,000	−.43	.11	.65***	−.39	.16	.67*
43,000+	−.24	.12	.79*	−.42	.17	.66*
Missing	−.11	.09	.89	−.19	.13	.83
Educational status:						
< high school	. . .	. . .	1.00	. . .	. . .	1.00
High school graduate	−.25	.08	.78**	−.38	.11	.69***
Some college	−.34	.10	.71***	−.42	.14	.66**
College graduate	−.38	.11	.68***	−.66	.17	.51***
Marital status:						
Married	. . .	. . .	1.00	. . .	. . .	1.00
Never married	−.28	.08	.76***	−.29	.11	.75**
Other[a]	−.12	.09	.89	−.23	.14	.80
1988 residence:						
Rural	. . .	. . .	1.00	. . .	. . .	1.00
Urban	−.03	.08	.97	−.02	.12	1.02
1984 use recency:						
Use in last year	−.20	.07	.81**	−.11	.11	.90
No use in last year	. . .	. . .	1.00	. . .	. . .	1.00
1984 lifetime use frequency	.40	.03	1.49***	−.54	.05	.58***

NOTE.—Model χ^2 (df = 23) for use reduction = 435.23***; model χ^2 (df = 23) for use denial = 405.11***.

[a] Divorced, widowed, or separated.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

was positively associated with use reduction and negatively associated with use denial. This reflects the trends observed in table 2. It also reflects the possible operation of methodological artifacts such as ceiling and floor effects (e.g., see Cook and Campbell 1979). Those in the highest use categories had more opportunity for reduction in use because of the way in which the use response categories were structured. Those in the lowest use category had nowhere else to go but to the nonuse category to become use deniers.

Cocaine use reduction and denial. Significant effects for use reduction and denial were detected for interview mode, race, education, marital status, and interviewer familiarity (see table 6). Compared with respondents interviewed alone face-to-face, those interviewed by telephone were significantly more likely to report less use and to deny use in 1988. Compared with white/other respondents, black respondents were significantly more likely to deny and report less cocaine use at follow-up. Hispanic respondents were also more likely than white/other respondents to deny cocaine use. Those with higher education levels (i.e., with at least some college) were less likely to report reduced use or to deny use than those who failed to finish high school. Compared with married respondents, those who were divorced, widowed, or separated or those who were never married were less likely to deny and report less cocaine use. Respondents who were interviewed by the 1988 interviewer on more than one occasion were also more likely than others to deny and report less use of cocaine in 1988. Respondent residence was associated with reduced use of cocaine but not with use denial. Residents of urban locations were less likely to report reduced use than residents of rural locations. As with marijuana reports, higher cocaine use frequency in 1984 was associated with increased rates of reduced use and lower rates of use denial. Use recency, income, gender, and age were not associated with cocaine use reduction or denial.[4]

Some overall model similarities should be underscored. The most consistent overall correlate of use reduction and denial, regardless of substance, was race. In particular, black respondents consistently underreported marijuana and cocaine use compared with white/other respondents. This effect is reinforced by the finding that odds ratios

---

4. Upon completion of these analyses, we were informed by a staff member at the National Opinion Research Center that a small percentage of the interviewers were assigned new identification numbers over the course of the study. This has the obvious effect of decreasing the precision of our estimates of "interviewer familiarity" effects. We reestimated all of the regression models eliminating the potentially problematic indicator of interviewer familiarity. Irrespective of the dependent variable being predicted, none of the parameter estimates or standard errors for the effects of other variables were substantively affected by our elimination of the interviewer familiarity indicator.

were consistently highest for estimates pertaining to black versus white/other comparisons. Black respondents consistently had over twice the odds of underreporting compared with whites/others—even after controlling for a range of other important correlates of use reduction. The strongest risk differences were observed for cocaine use denial; blacks had over three times the odds of denial compared with whites/others. Other consistent differences were observed for interview conditions (respondents interviewed by telephone always had an increased risk for underreporting), education (respondents with higher levels of education always had a decreased risk for underreporting), and marital status (married respondents always had an increased risk for underreporting). Region of respondent residence and interviewer familiarity were only salient in models predicting cocaine use denial. Respondent income was only salient in models predicting marijuana underreporting.

Following up on the significant findings with respect to race and on previous research comparing mode effects (Aquilino 1992), we constructed a series of models evaluating interactions between race/ethnicity and interview mode. Out of four models, only one such interaction approached significance—the interaction between Hispanic ethnicity and responses on a self-administered questionnaire in the model predicting cocaine use denial. The coefficient for this interaction suggested that Hispanic respondents who were interviewed with a self-administered questionnaire were significantly more likely to deny cocaine use than Hispanic respondents who were interviewed face-to-face. The relative underreporting for black respondents persisted, irrespective of interview condition.

Discussion

A high rate of underreporting was evident for both marijuana and cocaine, with rates of use reduction observed for just under one-third of all marijuana users and just over one-third of all cocaine users. If denial is the criterion for evaluating underreporting, it appears that the harder substance, cocaine, is more affected by this phenomenon. Outright denial of use at follow-up was nearly twice as prevalent among cocaine users as it was among marijuana users.

We have used the term "underreporting" to describe logically inconsistent responses in 1988 observed in this study. One major limitation of this study is the lack of an external validating criterion for 1984 responses. Our review of prior research on reinterview effects suggests that initial responses are likely to have more validity than those provided at follow-up; it also suggests that reinterviews tend to result in

Table 6. Logistic Regressions: Predicting Use Reduction and Denial of Cocaine Use (N = 1,505)

Variable	Use Reduction			Use Denial		
	b	Standard Error	Odds Ratio	b	Standard Error	Odds Ratio
1988 interview mode:						
Face-to-face/alone	. . .	. . .	1.00	. . .	. . .	1.00
Face-to-face/someone present	−.23	.28	.79	.39	.31	1.48
Self-report	−.20	.12	.82	−.09	.15	.91
Telephone	.68	.24	1.98**	1.17	.26	3.24***
1988 interviewer familiarity:						
One contact	. . .	. . .	1.00	. . .	. . .	1.00
> 1 contact	.30	.13	1.35*	.36	.15	1.43*
Race:						
White	. . .	. . .	1.00	. . .	. . .	1.00
Black	.79	.16	2.20***	1.20	.18	3.32***
Hispanic	.23	.17	1.25	.51	.20	1.67**
Age (years):						
24–25	. . .	. . .	1.00	. . .	. . .	1.00
26–27	−.06	.16	.94	−.07	.19	.93
28–29	−.30	.17	.74	−.19	.20	.83
30–32	−.12	.18	.89	−.22	.22	.80
Sex:						
Male	. . .	. . .	1.00	. . .	. . .	1.00
Female	.08	.12	1.08	.09	.14	1.09

Income:						
< 12,000	. . .	. . .	1.00	. . .	. . .	1.00
12,001–20,000	.12	.20	1.12	.16	.24	1.16
20,001–30,000	−.25	.20	.78	−.14	.25	.87
30,001–43,000	.01	.20	1.00	.13	.25	1.14
43,000 +	−.26	.21	.77	−.46	.28	.63
Missing	−.24	.19	.79	.04	.23	1.05
Educational status:						
< high school	. . .	. . .	1.00	. . .	. . .	1.00
High school graduate	−.32	.17	.73	−.33	.19	.72
Some college	−.39	.19	.68*	−.22	.22	.81
College graduate	−.62	.21	.54**	−.54	.26	.58*
Marital status:						
Married	. . .	. . .	1.00	. . .	. . .	1.00
Never married	−.49	.14	.61***	−.65	.17	.52***
Other[a]	−.39	.18	.67*	−.69	.23	.50***
1988 residence:						
Rural	. . .	. . .	1.00	. . .	. . .	1.00
Urban	−.34	.17	.71*	−.38	.20	.69
1984 use recency:						
Use in last year	−.18	.12	.83	−.16	.1[illegible]	.85
No use in last year	. . .	. . .	1.00	. . .	. . .	1.00
1984 lifetime use frequency	.73	.06	2.08***	−.25	.07	.78**

NOTE.—Model χ^2 (*df* = 23) for use reduction = 131.41***; model χ^2 (*df* = 23) for use denial = 242.30***.

[a] Divorced, widowed, or separated.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

reduced reports of behaviors (or symptoms) among those originally reporting them. There are, however, alternative explanations for logical inconsistencies observed in 1988 which may not reflect retest artifact. The validity of original reports may be threatened by errors in the data (erroneous coding, transcription, or data processing) or by respondent exaggeration. Deniers may reinterpret the question regarding use upon reinterview; experimental or low-level use may be honestly redefined as "nonuse" by certain respondents.[5]

Prior studies evaluating the effects of interview mode on underreporting in substance use surveys suffer from a similar problem. Researchers have assumed that within any given study, significantly lower rates of disclosure in one mode compared with another (usually telephone surveys compared with face-to-face interviews) reflect underreporting. Indeed, absent externally validating measures of substance use, an argument could be made that higher rates found in the face-to-face interviews reflect overreporting. Future research on underreporting in nationally representative samples needs to include externally validated measures of use (such as urine or hair samples, treatment records, or collateral reports by friends and relatives) against which respondent reports can be compared.

Another possible limitation of this study concerns its generalizability beyond the years covered by the two waves of interviews. Our analyses focus on reporting patterns observed during a period characterized by potent antidrug norms. Underreporting may be a reflection of the general antidrug sentiments held by society during the 1980s. To the extent that shifts away from such attitudes occur in other survey periods, the problem of underreporting may diminish in importance.

Rubio-Stipec et al. (1992) discuss three types of response errors in lifetime assessment of behavior using follow-up interviews. The present study focused on only one type of response error—failure to recall behavior reported earlier. Additional analyses suggest that one other type of error also occurs with relatively high prevalence within this survey: many subjects who initially denied substance use admit to use upon follow-up, with "onset" of use occurring before the 1984 interview. We are currently exploring the implications and correlates of this type of error. These additional analyses require reliance on the dating of behavior in a way that our present study of retest artifact does not; the precise dating of behaviors is notoriously inaccurate (see Jobe and Mingay 1991; Means and Loftus 1991).

These analyses suggested that there was differential underreporting by interview mode. An important limitation of the present study should be pointed out in connection with this finding. Although an attempt was made to randomly assign face-to-face and self-reported drug use interviews, our indicator of interview mode does not reflect a randomly assigned variable. In particular, telephone interviews and interviews in the presence of other people were not randomly assigned. Possible selection biases in mode assignment could have influenced our findings. Nevertheless, the influence of such biases were limited by our inclusion of demographic control variables (variables that should be associated with mode assignment) in multiple logistic regression models.

---

[5] These possibilities were suggested by an anonymous reviewer.

The validity of this quasi-experimental approach in the assessment of interview mode effects is supported by the consistency of our findings with previous studies. Consistent with previous studies cited above, this study suggests that telephone interviews lead to significant underreporting of substance use compared with face-to-face interviews or self-administered questionnaires. The lack of difference between underreporting for face-to-face interviews and self-administered questionnaires is not particularly surprising in light of the overall format for the NLSY interview. The drug use section of the survey was a separate section; all respondents had an interviewer with whom they interacted prior to the completion of this section. The presence of an interviewer and the repeated contact with study personnel over multiple interviews may have outweighed the additional sense of privacy gained through self-administration. Our interview mode comparisons provide evidence that telephone interviews are an ineffective method for ascertaining substance use history in longitudinal studies. The conclusions drawn by Johnson, Hougland, and Moore (1991) are relevant in this context. These researchers suggested that telephone interviews fail to provide respondents with "obvious cues to interviewer identity and the opportunities for nonverbal interaction routinely available in a face-to-face situation" (p 181). Such cues may be necessary to encourage self-revelation, especially under circumstances in which respondent anonymity is threatened (such as longitudinal surveys).

Analyses suggest that, irrespective of drug use survey mode, repeated contact with the same interviewer can have an impact on responses with respect to lifetime use of drugs. Consistent with the suggestions made by Mensch and Kandel (1988), interviewer familiarity does lead to underreporting. Note, however, that the extent of underreporting must be narrowly circumscribed. Interviewer familiarity was not associated with a decline in reports of marijuana use. Interviewer familiarity *was*, however, associated with higher rates of cocaine use underreporting.

In general, subgroup differences remained even after mode was controlled for in the multivariate logistic regression model. Consistent with prior literature suggesting differential sensitivities to interview mode by race, it appeared that race was the single most consistent correlate of underreporting behavior. Hispanic respondents consistently had higher odds of underreporting compared with white respondents. Nevertheless, black versus white/other differences were striking in our analyses. Black respondents had at least twice the odds of underreporting compared with white/other respondents for every measure of underreporting.

Differences in rates of underreporting by race/ethnicity need to be examined further, especially in light of recent epidemiological studies showing lower rates of reported use among minority respondents. In the *Monitoring the Future* study of high school seniors, Johnston, Bachman, and O'Malley (1986) found that, compared with white respondents, black respondents were much more likely to leave responses about drug use blank. Additionally, in supplemental questions, black respondents indicated that if they used drugs, they would not have been willing to report it. Recent waves of the *Monitoring the Future* study summarizing the trends in prevalence across race/ethnic groups suggest relatively low levels of substance use prevalence (based on a 1-year reporting frame) for nonwhite youths compared with white youths (Bachman et al. 1991). These differences were more pronounced for cocaine than they were for marijuana. Similar race/ethnicity differences have been found in other school-based surveys of adolescents (see Oetting and Beauvais [1990] for a recent review). Research on samples of juvenile arrestees (Fendrich and Xu 1994) suggests that race/ethnicity differences in underreporting are not limited to school-based surveys. These prior results, when evaluated in the context of our findings with a longitudinally followed sample of adults, suggest that prevalence differences across race/ethnicity groups may be due in part to differences in underreporting and not entirely to differences in actual use.

In order to improve the validity of drug use reporting in survey research, the precise reasons for differential rates of underreporting among minority respondents need to be explored in greater detail. Our analyses suggest that these differences persist even after a number of possible confounding covariates are controlled for (income, education, residence, and use level at the initial interview). Thus, an explanation for persistent differences might focus on aspects of the interview process itself. A significant race-by-interview-condition interaction between Hispanic respondents and self-reports may indicate that some race/ethnicity differences are a function of language comprehension. Self-report measures written in English may prove challenging to some minority respondents whose native language is not English. Nevertheless, this explanation is of limited relevance to the persistent black versus white/other differences uncovered here.

One promising line of inquiry might investigate race/ethnicity differences between minority respondents and interviewers (Campbell 1981; Columbotos, Elinson, and Lowenstein 1968; Weeks and Moore 1981). Unfortunately, the data set did not facilitate our analysis of this issue. It should, however, be stressed that the underreporting found in this study characterizes data obtained from a number of interview settings. Furthermore, race differences between interviewer and subject are less salient in telephone interviews, the one setting found to show considerable race/ethnicity differences in underreporting in the past.

Previously, in their explanation of race differences in marijuana underreporting, Mensch and Kandel (1988) cited prior comments by Johnston, Bachman, and O'Malley (1984) with respect to black respondents' lack of trust in the research process. This lack of trust may be grounded in previous experience and knowledge about the consequences of disclosing information about sensitive behavior. The reasons associated with underreporting for race may be relevant to a discussion of underreporting for other respondents in which perceived inequalities or social status concerns may be salient. Status inequalities perceived by respondents with lower levels of education or income may result in an increased sense of vulnerability and a general reluctance to disclose information about illicit, socially unacceptable behavior.

Recent research has taken cognitive approaches to understand the validity of behavioral reports and to structure interviews to facilitate more valid reporting (see Jobe and Mingay 1991; Means and Loftus 1991). This work emphasizes the role of memory and comprehension in the validity of self-reports about behavior. The influence of low frequency use on underreporting and denial for both marijuana and cocaine may suggest that fading memory for rare, distant events may influence behavioral reports. If this were the case, however, one would expect use recency in 1984 to have a consistent negative association with underreporting; we did not find this to be the case. Additionally, the logistic regression models demonstrated that the influence of other variables on underreporting persisted even after use level at the initial interview was controlled for. These models suggest that underreporting may be more contingent on motivation than it is on memory.

References

Anglin, M. Douglas, Yih-Ing Hser, and Chih-Ping Chou. 1993. "Reliability and Validity of Retrospective Behavioral Self-Report by Narcotic Addicts." *Evaluation Review* 17:91–108.

Aquilino, William S. 1992. "Telephone versus Face-to-face Interviewing for Household Drug Use Surveys." *International Journal of the Addictions* 27:71–91.

Aquilino, William S., and Leonard A. LoSciuto. 1990. "Interview Mode Effects in Drug Use Surveys." *Public Opinion Quarterly* 54:362–95.

Bachman, Jerald G., John M. Wallace, Patrick M. O'Malley, Lloyd D. Johnston, Candace L. Kurth, and Harold W. Neighbors. 1991. "Racial/Ethnic Differences in Smoking, Drinking, and Illicit Drug Use among American High School Seniors, 1976–1989." *American Journal of Public Health* 81:372–77.

Bromet, Evelyn, Leslie Dunn, Melanie Connell, Mary Dew, and Herbert Schulbert. 1986. "Long-Term Reliability of Diagnosing Lifetime Major Depression in a Community Sample." *Archives of General Psychiatry* 43:435–40.

Campbell, Bruce A. 1981. "Race-of-Interviewer Effects among Southern Adolescents." *Public Opinion Quarterly* 45:231–44.

Center for Human Resource Research. 1991. *NLS Handbook 1991*. Columbus: Ohio State University, Center for Human Resource Research.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

Collins, Linda, John Graham, William B. Hansen, and C. Anderson Johnson. 1985. "Agreement between Retrospective Accounts of Substance Use and Earlier Reported Substance Use." *Applied Psychological Measurement* 9:301–9.

Columbotos, John, Jack Elinson, and Regina Lowenstein. 1968. "Effect of Interviewers' Sex on Interview Responses." *Public Health Reports* 83:685–90.

Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.

Eaton, W. W., M. Kramer, J. C. Anthony, A. Dryman, S. Shapiro, and B. Z. Locke. 1989. "The Incidence of Specific DIS/DSM-III Mental Disorders: Data from the NIMH Epidemiological Catchment Area Program." *Acta Psychiatrica Scandanavia* 79:163–68.

Fendrich, Michael, Paul Goldstein, Chaim Tarshish, and Patricia Bellucci. 1992. "Longitudinal Measurement of Substance Use in Ethnographic Samples." *Journal of Community Psychology* 20:236–342.

Fendrich, Michael, and Virginia Warner. In press. "Symptom and Substance Use Recall over Two Years for Offspring at High and Low Risk for Depression." *Journal of Abnormal Child Psychology*.

Fendrich, Michael, and Yanchun Xu. 1994. "The Validity of Drug Use Reports from Juvenile Arrestees." *International Journal of the Addictions*, vol. 29 (in press).

Gfroerer, Joseph C., and Arthur L. Hughes. 1991. "The Feasibility of Collecting Drug Abuse Data by Telephone." *Public Health Reports* 106:384–93.

———. 1992. "Collecting Data on Illicit Drug Use by Phone." *Survey Measurement of Drug Use: Methodological Studies*, ed. Charles F. Turner, Judith T. Lessler, and Joseph C. Gfroerer, pp. 277–95. Rockville, MD: U.S. Department of Health and Human Services, Alcohol, Drug Abuse, and Mental Health Administration.

Henderson, Scott, D. G. Byrne, and Paul Duncan-Jones. 1981. *Neurosis and the Social Environment*. New York: Academic Press.

Hosmer, David W., and Stanley Lemeshow. 1989. *Applied Logistic Regression*. New York: Wiley.

Jobe, Jared B., and David J. Mingay. 1991. "Cognition and Survey Measurement: History and Overview." *Applied Cognitive Psychology* 5:175–92.

Johnson, Timothy P., James G. Hougland, and Richard R. Clayton. 1989. "Obtaining Reports of Sensitive Behavior: A Comparison of Substance Use Reports from Telephone and Face-to-Face Interviews." *Social Science Quarterly* 70:174–83.

Johnson, Timothy P., James G. Hougland, and Robert W. Moore. 1991. "Sex Differences in Reporting Sensitive Behavior: A Comparison of Interview Methods." *Sex Roles* 24:669–80.

Johnston, Lloyd D., Jerald G. Bachman, and Patrick M. O'Malley. 1984. *Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1983*. Ann Arbor: University of Michigan, Institute for Social Research.

———. 1986. *Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1985*. Ann Arbor: University of Michigan, Institute for Social Research.

Johnston, Lloyd D., Patrick M. O'Malley, and Jerald G. Bachman. 1988. *Illicit Drug Use, Smoking, and Drinking by America's High School Students, College Students, and Young Adults, 1985–1987*. Rockville, MD: National Institute on Drug Abuse.

Jones, Elise F., and Jacqueline D. Darroch-Forrest. 1992. "Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988." *Demography* 29:113–26.

Jorm, A. F., P. Duncan-Jones, and R. Scott. 1989. "An Analysis of the Retest Artifact in Longitudinal Studies of Psychiatric Symptoms and Personality." *Psychological Medicine* 19:487–93.

Means, Barbara, and Elizabeth Loftus. 1991. "When Personal History Repeats Itself: Decomposing Memories for Recurring Events." *Applied Cognitive Psychology* 5:297–318.

Mensch, Barbara S., and Denise B. Kandel. 1988. "Underreporting of Substance Use in a National Longitudinal Youth Cohort: Individual and Interviewer Effects." *Public Opinion Quarterly* 52:100–124.

Newcomb, Michael D., and Peter M. Bentler. 1988. *Consequences of Adolescent Drug Use: Impact on the Lives of Young Adults*. Newbury Park, CA: Sage.

Oetting, Eugene R., and Fred Beauvais. 1990. "Adolescent Drug Use: Findings of National and Local Surveys." *Journal of Consulting and Clinical Psychology* 58:385–94.

Rice, John B., Jean Endicott, Mary Ann Knesevich, and Nanette Rochberg. 1987. "The Estimation of Diagnostic Sensitivity Using Stability Data: An Application to Major Depressive Disorder." *Journal of Psychiatric Research* 21:337–45.

Rubio-Stipec, Maritz, Daniel H. Freeman, Lee Robins, Patrick Shrout, Glorisa Canino, and Milagros Bravo. 1992. "Response Error and the Estimation of Lifetime Prevalence and Incidence of Alcoholism: Experience in a Community Survey." *International Journal of Methods in Psychiatric Research* 2:501–8.

Schober, Susan, M. Fe Caces, Michael Pergamit, and Laura Branden. 1992. "Effect of Mode of Administration on Reporting of Drug Use in the National Longitudinal Survey." In *Survey Measurement of Drug Use: Methodological Studies*, ed. Charles F. Turner, Judith T. Lessler, and Joseph C. Gfroerer, pp. 267–76. Rockville, MD: U.S. Department of Health and Human Services, Alcohol, Drug Abuse, and Mental Health Administration.

Sudman, Seymour, and Norman M. Bradburn. 1974. *Response Effects in Surveys*. Chicago: Aldine.

Turner, Charles, Judith Lessler, and James Devore. 1992. "The Effects of Mode of Administration and Wording on Reporting of Drug Use." In *Survey Measurement of Drug Use: Methodological Studies*, ed. Charles F. Turner, Judith T. Lessler, and Joseph C. Gfroerer, pp. 177–219. Rockville, MD: U.S. Department of Health and Human Services, Alcohol, Drug Abuse, and Mental Health Administration.

Weeks, Michael F., and R. Paul Moore. 1981. "Ethnicity-of-Interviewer Effects on Ethnic Respondents." *Public Opinion Quarterly* 45:245–49.

Yamaguchi, Kazuo, and Denise B. Kandel. 1984. "Patterns of Drug Use from Adolescence to Young Adulthood. II. Sequences of Progression." *American Journal of Public Health* 74:668–72.

THE POLLS—POLL TRENDS
THE RISE OF THE NEW MEDIA

WILLIAM G. MAYER

One of the most talked about trends in media and communications studies over the past decade has been the rise of the so-called new media.[1] The phrase "new media" is, of course, a relative one. At one time, the printing press would have been considered a new medium; the same could have been said about radio in the 1920s and television in the 1950s. In its most recent incarnation, however, the phrase has been widely adopted to refer to a series of scientific and technological innovations, most of which achieved critical mass in the early 1970s, that led to the development of a considerable number of new methods for creating, transmitting, and storing information and to the large-scale transformation of many of the more traditional media. Among the technologies usually included under the "new media" rubric are personal computers, videocassette recorders (VCRs), videotex and teletext, fax machines, car phones, satellites and satellite dishes, remote control devices, videodiscs, compact discs, modems, telephone answering machines, electronic mail, cable TV, and interactive television. Though academic researchers were initially slow to pick up on these developments, a considerable body of writing has recently begun to accumulate, examining the characteristics of the new media and the impacts they are likely to have on the way Americans work, learn, conduct their politics, and spend their leisure time (see, among others, Abramson, Arterton, and Orren 1988; Arterton 1987; Barber 1984; Gilder 1992; Naisbitt 1982; Neuman 1991; Williams 1982).

Until now, most of the literature on the new media has necessarily been speculative. That is to say, its focus has been on how American media habits will *change* and what our politics and society will look like *in the future*, after these new technologies have had a chance to

WILLIAM G. MAYER is assistant professor of political science at Northeastern University.
The author wishes to thank Lois Timms-Ferrara, Rosemary Kryk, Michael Dukakis, and Robert Y. Shapiro for their generous assistance in the preparation of this article.

---

[1] For a good overview of the trends and possibilities in research on the new media, see Rice 1984, esp. chaps. 1–4.

Public Opinion Quarterly Volume 58:124–146 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5801-0004$02.50

be more widely adopted and integrated into American life.[2] But we are rapidly approaching the time when such speculation will no longer be necessary: when the new media will no longer be so new and exotic, and communications scholars will be able to study the way these technologies are actually being used (or not used) by the American mass public.

In an earlier article, I presented survey data on the trends in public usage of such "traditional" media as newspapers, television, and radio (see Mayer 1993). This article carries the story forward by examining the available data on how the American people have reacted to the new media.

Since these media are so new, so are the survey questions presented here. With one exception, none of the questions reported in this article was asked before 1980. Even so, enough data have accumulated to permit some early conclusions about the public's response to this new "smorgasbord" of media alternatives: which of the new technologies have been received most enthusiastically by the mass public and, in a few cases, how and how often these new media are being used.

The data are richest (and probably most reliable) in assessing public ownership of the various new media products. On this score, the data provide further evidence that, in some cases, new media technologies can be adopted by the mass public at a remarkably rapid rate (see Mayer 1993). As recently as 1980, for example, only 2 percent of the population owned VCRs; by 1991, 73 percent owned them. The number subscribing to cable television jumped from 15 percent in 1974 to 62 percent in 1990. Where only 3 percent of U.S. households had a telephone answering machine in 1981, by 1992 45 percent owned one.

Yet, not all the new technologies have been eagerly snatched up by the American public. As a number of scholars have emphasized, the way a nation communicates is dependent not only on what is technologically possible but also on what the average person is willing and able to use and pay for (see, esp., Neuman 1991). As of the early 1990s, there were clear gradations in the extent to which the new media technologies had penetrated the mass public:

Over 70 percent of American households owned a VCR and had a remote control device for their television set.
About 60 percent subscribed to cable TV.
About 45 percent owned a telephone answering machine.

[2] It is worth adding that, even under the best circumstances, forecasting the growth of a new media product or service is a difficult and complex task, and that most forecasts are not very accurate beyond a very short-term time horizon. See esp. Lazarus and McKnight 1984.

Personal computers and compact disc players were owned by about 20–30 percent of all households.

Relatively few Americans, however—generally on the order of about 2 or 3 percent—owned a fax machine, a satellite dish, a car phone, a modem, or a laptop computer.

As the data in the second section below indicate, however, ownership is not the only measure of how many Americans have been touched by the various new media technologies. Many people who do not personally own one of the new media products may nevertheless have access to one. Where only about 1 percent of Americans own a fax machine, for example, 27 percent say they have access to one at work. In 1986, when only about 15 percent owned a personal computer, about 30 percent said they worked at a place where computers were used.

The data are considerably thinner in assessing usage rates and patterns for the new media. The problem, of course, is that if only 1 percent of Americans say they own a fax machine, then it is not very useful, given the size of conventional sample surveys, to ask these fax owners a series of follow-up questions about how and how often they use their machines. As a result, only a few of the new media—VCRs, cable television, personal computers, and remote control devices, in particular—have been the subject of more extensive examination.

One important conclusion we can draw from these usage data is that as the new media achieve a greater penetration, they are likely to supplement and modify existing media, but not to replace them entirely.[3] A good example of this is a Roper question that asked respondents about the kinds of things they had "done for entertainment . . . during the past week." The percentage saying that they had "played a videotape or disc" rose from 6 percent in 1982 to 29 percent in 1990. But the same 1990 survey showed that 78 percent of Americans had watched a movie on TV during the past week, 56 percent had read a magazine, and 44 percent had read a book. In a 1991 Roper survey, 29 percent of the sample said they had rented a videotape during the past week, and 25 percent had watched a movie on a pay-cable channel such as HBO or Showtime. But 63 percent had watched a movie on regular broadcast TV, and 12 percent had gone to a movie theater. Many of the questions show large and rapid changes in American media usage patterns—but a fascination with change should not obscure the substantial continuity in media-related behavior.

There is also some indication that as the percentage of Americans who own a new media technology increases, that technology is used less often and in less innovative ways than might have been predicted on the basis of reported behavior by early owners. For example, as VCRs have spread from 2 percent to 73 percent of American households, there is evidence that VCR owners tape fewer TV shows per week, rent fewer movies, and are less likely to eliminate commercials as they replay taped material.

---

[3] The previous literature on the new media has often suggested that this would be the case. See, e.g., Abramson, Arterton, and Orren 1988; Neuman 1991; and Rice 1984, esp. chap. 12.

Most of the questions reported here are drawn from surveys conducted by the Roper Organization, the results of which are published in *Roper Reports*. All other data are based on reports or news releases put out by the survey organization listed, which are available in the archives of the Roper Center for Public Opinion Research, Storrs, CT. "Don't know" responses have been included in percentaging all results, but are only reported when they equal more than 1 percent of the total.

Appendix
Sources and Abbreviations

AIPO: American Institute for Public Opinion (Gallup)
Barna: Barna Research Group
Black: Gordon S. Black/*U.S.A. Today*
CBS: CBS News
Roper: Roper Organization

References

Abramson, Jeffrey B., F. Christopher Arterton, and Gary R. Orren. 1988. *The Electronic Commonwealth*. New York: Basic.

Arterton, F. Christopher. 1987. *Teledemocracy: Can Technology Protect Democracy?* Newbury Park, CA: Sage.

Barber, Benjamin J. 1984. *Strong Democracy: Participatory Politics for a New Age*. Berkeley and Los Angeles: University of California Press.

Gilder, George. 1992. *Life after Television*. New York: Norton.

Lazarus, William, and Lee McKnight. 1984. "The Design and Forecasting of the New Media." Discussion paper, Massachusetts Institute of Technology, Research Program on Communications Policy.

Mayer, William G. 1993. "Poll Trends: Trends in Media Usage." *Public Opinion Quarterly* 57:593–611.

Naisbitt, John. 1982. *Megatrends*. New York: Warner.

Neuman, W. Russell. 1991. *The Future of the Mass Audience*. Cambridge: Cambridge University Press.

Rice, Ronald E., ed. 1984. *The New Media: Communication, Research, and Technology*. Beverly Hills, CA: Sage.

Williams, Frederick. 1982. *The Communications Revolution*. Beverly Hills, CA: Sage.

Ownership

Do you happen to have a video cassette recorder—a VCR—in your household or not?

		Roper	
Survey Date	*Yes (%)*	*No (%)*	*N*
July 1980[a]	2	98	2,005
October–November 1983[b]	6	93	1,000
October–November 1984[c]	16	84	2,000
February 1985[c]	19	81	2,000
September 1985[c]	23	77	1,996
September 1986[c]	37	62	1,997
October 1986[c]	43	57	1,984
September 1987	48	52	1,990
March 1988	58	42	1,993
September 1988	61	39	1,996
May 1989	67	33	1,986
September 1989	65	35	1,973
April 1990	71	29	1,978
September 1991	73	26	1,973

[a] Do you happen to own a video tape recorder?

[b] Do you happen to own a video cassette recorder?

[c] Do you happen to own a video cassette recorder with which you can tape shows that appear on TV, or not?

Here's a list of some utensils and appliances people have in their homes. [Card shown respondent.] Would you look down it and call off the ones, if any, that you have in your home right now?

		Roper	
	July 1981 (%)	*July 1986 (%)*	*July 1989 (%)*
Compact disc player	. . .	6	11
Telephone answering machine	3	13	25
Home computer	1	14	15
Modem	. . .	. . .	4
Home video games	9	21	23
Video disc player	. . .	. . .	14
N	(2,000)	(1,991)	(1,979)

Here is a list of some different products. [Card shown respondent.] Would you read down that list and call off those that you now own?

	March 1982 (%)	March 1983 (%)	Roper March 1984 (%)	March 1986 (%)	February 1987 (%)
Color TV set	83	84	87	90	. . .
Black and white TV set	54	53	52	41	. . .
Stereo equipment	67	70	73	69	. . .
Video tape recorder[a]	8	10	21	35	45
Video disc player	2	3	4	3	5
Electronic TV game	12	18	24	. . .	22
Home computer	2	4	9	11	14
Camera for still pictures[b]	. . .	. . .	70	69	73
Movie camera	. . .	. . .	18	12	14
Video tape camera[c]	. . .	. . .	2	4	6
Compact audio disc player	. . .	. . .	. . .	2	. . .
N	(2,000)	(2,000)	(2,000)	(1,993)	(998)

[a] In 1984 and 1986, "video cassette recorder"; in 1987, "video tape recorder/player."

[b] In 1987, "camera for still photographs."

[c] In 1987, "video camera."

Here is a list of some different products. [Card shown respondent.] Would you read down that list and call off those that you now have in your household?

	February 1987 (%)	Roper February 1990 (%)	February 1992 (%)
TV set	98	98	97
Record player/turntable[a]	75	66	59
Audio tape recorder/player	55	56	52
Video tape recorder/player[b]	46	64	68
Video disc player	5	9	11
Electronic TV game	22	26	32
Home computer	14	18	20
Walkman type radio or tape player	37	44	44
Camera for still photographs	72	71	67
Movie camera	14	15	. . .
Video camera (camcorder)[c]	6	12	15
Compact disc player	. . .	13	21
N	(998)	(1,977)	(1,997)

[a] In 1987, "record player."

[b] In 1992, "video cassette recorder/player."

[c] In 1987, "video camera."

Here is a list of some different products. [Card shown respondent.] Would you read down that list and for each one tell me whether that is something you now own or not?

	Roper (May–June 1984; N = 2,000)	
	Own (%)	*Don't Own (%)*
A beeper for paging	5	94
A cordless telephone	12	88
A battery operated transistor radio	72	28
A pocket size TV set with a tiny 2-inch screen	1	98
A telephone in your car	1	99

Which of these things [card shown respondent] do you have in your household? Please just call off the letters opposite the items you now have. Which of the others, if any, would you like to have? Again, please call off the letters.

		Roper	
	Now Have (%)	*Would Like to Have (%)*	*N*
A satellite "dish" to receive cable TV:			
July 1984	3	20	2,000
December 1986	3	16	1,997
A remote control to change TV stations so you don't have to go to the set to do it:			
July 1984	28	17	2,000
December 1986	46	16	1,997

Here is a list of some different kinds of things some people own. [Card shown respondent.] Would you read down that list, and for each one tell me whether it is something now owned in your household, or something you definitely plan to buy within the next few months, or something you may buy sometime, or something you're not at all interested in owning?

Roper

	Own Now (%)	*Plan to Buy (%)*	*May Buy (%)*	*Not Interested (%)*	*Don't Know (%)*	*N*
A home computer:						
December 1982	3	6	22	67	3	2,000
December 1985	11	6	25	56	2	1,980
December 1988	14	7	24	53	3	1,974
A video cassette recorder (VCR):						
December 1982	6	4	18	69	3	2,000
December 1985	31	11	21	35	1	1,980
December 1988	64	7	10	18	1	1,974
A compact audio disc player (CD):						
December 1985	4	4	14	74	5	1,980
December 1988	11	6	18	62	2	1,974
A telephone answering machine:						
December 1985	13	5	14	66	2	1,980
December 1988	23	6	13	56	2	1,974
Fax or facsimile machine:						
December 1988	0	1	4	89	5	1,974
A telephone in your car:						
December 1988	1	1	8	87	3	1,974
A video camcorder:						
December 1988	5	8	22	61	3	1,974
A satellite TV dish:						
December 1988	2	2	11	82	3	1,974

[Asked of respondents who owned a home computer:] You say you have a home computer. Do you have a modem for that computer or not?

Roper

	Yes (%)	*No/ Don't Know (%)*	*Don't Own a Computer (%)*	*N*
February 1987	3	11	86	1,996

As technology has improved over the years, telephones have become increasingly sophisticated and able to perform new functions. Here is a list of some of the things telephones can now do. [Card shown respondent.] Would you please read down that list and tell me whether you have any telephone at home that does that or don't you have any phone that does that?

	Roper	
	December 1987 (%)	*December 1992 (%)*
Touch-tone push-button dialing	68	79
Call-waiting that tells you if someone is calling you while you are talking on the phone	26	42
A speakerphone that allows you to talk without using the telephone receiver	8	17
A telephone answering machine that answers the telephone and records the message left by the caller after the beep	17	45
Automatic redial of the last number you called	38	57
Automatic one-button dialing of numbers you have programmed into the telephone's memory	21	36
A cordless phone that allows you to use the phone anywhere in the house or yard	18	32
A telephone on your car	1	5
A portable cellular phone that you can carry anywhere	. . .	4
A way to identify the phone number of people who call you (Caller I.D.)	. . .	3
A video telephone that allows you to speak and see the person on the other end of the line	. . .	1
A permanent phone number that stays with you regardless of where you live	. . .	9
N	(1,980)	(1,977)

NOTE.—Figures are the percentage who say they *have* the function or product listed.

Does the TV set you tend to watch most often have a remote control you can use to change from channel to channel—either through your TV set or your cable TV box?

			Roper	
	Yes (%)	*No (%)*	*Don't Know (%)*	*N*
July 1988	57	41	2	1,968

[Asked of respondents with a television in the house:] Do you have a remote control on at least one television set such that you can change channels from a distance or not?

	AIPO *Yes* *(%)*	*No* *(%)*	*Don't Own a Television* *(%)*	*N*
August 1990	79	20	2	1,241

Which, if any, of the following types of equipment does your household own?

	Barna (January–February 1991; N = 1,005) *(%)*
A VCR (video cassette recorder)	78
A cellular telephone for your car	7
A home fax machine	2
A home computer	26
A compact disc player	30
A satellite dish for TV (television) reception	6
A telephone answering machine	46

Here is a list of some other electronic products for the home. [Card shown respondent.] Which of these products have you ever heard of? Which of the listed items do you own?

	Roper (February 1992; N = 1,997)	
	Heard of *(%)*	*Own* *(%)*
DAT (Digital Audio Tape)	30	1
HDTV (High Definition TV)	28	1
A fax machine	79	2
Super VHS	33	3
VCR Plus	43	8
Electronic organizer/calendar	30	3
Hand-held portable TV/VCR unit	41	3
CD player that records	37	3
Pocket cellular telephone	43	1
Scientific calculator	32	7
Car telephone	85	5
Lap-top computer	55	3
Laser disc	49	1
Video phone	41	0
Interactive television (The Imagination Machine, CDTV, etc.)	13	0

Access

As far as you know, is cable TV available here in your neighborhood? [If yes:] Are you a subscriber to cable TV or not?

			Roper		
	January 1974 (%)	*January 1978 (%)*	*January 1980 (%)*	*January 1982 (%)*	*April 1982 (%)*
Yes, cable TV available	29	42	52	63	62
And subscribe	15	18	24	35	35
Do not subscribe	14	24	28	26	27
No, cable TV not available	58	45	40	32	33
Don't know	13	13	8	5	5
N	(1,997)	(2,002)	(2,005)	(2,000)	(2,000)

	December 1982 (%)	*December 1983 (%)*	*March 1984 (%)*	*April 1984 (%)*	*July 1984 (%)*
Yes, cable TV available	65	79	79	75	78
And subscribe	37	45	48	43	46
Do not subscribe	28	34	31	32	31
No, cable TV not available	31	17	18	23	20
Don't know	4	4	3	2	2
N	(2,000)	(2,000)	(2,000)	(2,000)	(2,000)

	September 1984 (%)	*December 1984 (%)*	*February 1985 (%)*	*April–May 1985 (%)*	*June 1985 (%)*
Yes, cable TV available	78	79	79	78	80
And subscribe	45	47	48	44	47
Do not subscribe	32	32	31	34	33
No, cable TV not available	20	19	18	19	18
Don't know	2	2	2	3	2
N	(2,000)	(1,977)	(2,000)	(1,988)	(2,000)

	July 1985 (%)	*August 1985 (%)*	*July 1988[a] (%)*
Yes, cable TV available	78	80	84
And subscribe	45	47	56
Do not subscribe	33	32	28
No, cable TV not available	20	18	13
Don't know	2	1	1
N	(1,997)	(1,996)	(1,968)

[a] In 1988, second question read, "Do you subscribe to cable TV or not?"

Do you subscribe to cable TV or not? [If yes:] In addition to your basic cable service, do you also pay extra to get special sports and entertainment channels such as HBO, Showtime, The Sports Channel, etc.?

		Roper	
	May 1988 (%)	*September 1989 (%)*	*January 1990 (%)*
Yes, subscribe	55	56	62
Yes, have extra channels	30	32	34
No, do not have extra channels	24	24	28
Do not subscribe	45	43	38
N	(1,962)	(1,973)	(1,994)

Do you work in a place where either computer terminals or word processors are used? [If yes:] Do you yourself work at a computer terminal or word processor almost all the time, or frequently, or occasionally, or never?

		Roper	
	April 1982 (%)	*April–May 1985 (%)*	*February 1986 (%)*
Yes, work in place where they are used	24	33	27
Almost all the time	. . .	. . .	4
Frequently	. . .	. . .	5
Occasionally	. . .	. . .	8
Never	. . .	. . .	10
No, do not work in a place where they are used	76	67	73
N	(2,000)	(1,988)	(1,993)

Here is a list of some types of clubs different people belong to. [Card shown respondent.] Would you read down that list and call off those, if any, that you happen to belong to?

	Roper *(April–May 1985; N = 1,988)* *(%)*
A computer network or electronics billboard club	1

Do you have access to a FAX machine—either at home or at work?

			CBS	
	Yes (%)	*No (%)*	*Don't Know/ No Answer (%)*	*N*
June 1989	27	71	2	1,225

Do you personally use any of these machines at a place where you work? [Card shown respondent.]

		Roper	
	January 1983 (%)	*January 1986 (%)*	*January 1990 (%)*
Word processor	5	11	14
Computer	13	19	27
Computer terminal	11	14	19
Electric typewriter	20	23	21
Copying machine	24	29	30
Fax machine	. . .	. . .	15
No/none/don't work	67	63	59
N	(2,000)	(2,000)	(1,994)

Usage Rates

Now here is a list of a number of different things. [Card shown respondent.] Would you read down that list and call off each one you personally have done in the last month?

		Roper	
	October 1981 (%)	*April 1982 (%)*	*October 1982 (%)*
Bought a compact disc record	. . .	. . .	. . .
Bought a regular record	. . .	11	20
Bought a pre-recorded music tape	. . .	7	11
Bought a blank audio tape	. . .	7	14
Bought a blank video cassette tape	. . .	1	. . .
Taped something on a video cassette recorder	5	. . .	5
Rented a movie to play on a video cassette player	. . .	. . .	. . .
Bought a movie to play on a video cassette player	. . .	. . .	. . .
N	(2,000)	(2,000)	(2,000)
	April 1984 (%)	*October–November 1984 (%)*	*October–November 1985 (%)*
Bought a compact disc record	. . .	. . .	. . .
Bought a regular record	16	15	13
Bought a pre-recorded music tape	11	13	15
Bought a blank audio tape	11	10	10
Bought a blank video cassette tape	5	8	12

	April 1984 (%)	*October–November 1984 (%)*	*October–November 1985 (%)*
Taped something on a video cassette recorder	. . .	. . .	19
Rented a movie to play on a video cassette player	. . .	10	20
Bought a movie to play on a video cassette player	. . .	2	3
N	(2,000)	(2,000)	(1,998)

	April–May 1986 (%)	*October 1986 (%)*	*April–May 1987 (%)*
Bought a compact disc record[a]	2	1	3
Bought a regular record[b]	15	10	12
Bought a pre-recorded music tape	13	13	15
Bought a blank audio tape	11	12	12
Bought a blank video cassette tape	12	15	18
Taped something on a video cassette recorder	. . .	25	. . .
Rented a movie to play on a video cassette player[c]	. . .	27	. . .
Bought a movie to play on a video cassette player[d]	. . .	4	. . .
N	(1,998)	(1,984)	(2,000)

	October 1987 (%)	*April 1990 (%)*	*April 1991 (%)*
Bought a compact disc record[a]	4	9	8
Bought a regular record[b]	10	8	6
Bought a pre-recorded music tape	18	19	12
Bought a blank audio tape	14	16	. . .
Bought a blank video cassette tape	18	18	. . .
Taped something on a video cassette recorder	28	. . .	. . .
Rented a movie to play on a video cassette player[c]	36	33	24
Bought a movie to play on a video cassette player[d]	7	11	7
N	(1,990)	(1,978)	(1,984)

[a] In 1990 and 1991, "bought a compact disc recording."
[b] In 1990 and 1991, "bought a record."
[c] In 1990 and 1991, "rented a pre-recorded video tape."
[d] In 1990 and 1991, "bought a pre-recorded video tape."

Now here is a list of different things. [Card shown respondent.] Would you read down that list and call off each one you personally have done in the last week, either at home or at work?

	Roper			
	September 1986 (%)	*September 1987 (%)*	*September 1988 (%)*	*September 1989 (%)*
Used a computer at work or at home	16	18	19	18
Taped something on a VCR	. . .	. . .	. . .	. . .
Rented a video tape	. . .	. . .	. . .	21
Ordered a pay-per-view movie by phone	. . .	. . .	. . .	. . .
Gone to the movies	. . .	12	. . .	. . .
Watched a movie on a pay-cable channel (HBO, Showtime, etc.)	. . .	. . .	. . .	. . .
Watched a movie on a basic cable channel (A&E, TNT, TBS, USA, etc.)	. . .	. . .	. . .	. . .
Watched a movie on regular broadcast TV (ABC, CBS, NBC, PBS, Fox, independents, etc.)	. . .	. . .	. . .	. . .
Took snapshots	. . .	. . .	. . .	. . .
Had picture taken by someone	. . .	. . .	. . .	. . .
Used a video camera	. . .	. . .	. . .	. . .
N	(1,997)	(1,990)	(1,996)	(1,973)

	September 1990 (%)	*March 1991 (%)*	*September 1991 (%)*
Used a computer at work or at home	20	. . .	17
Taped something on a VCR	. . .	. . .	23
Rented a video tape	. . .	. . .	29
Ordered a pay-per-view movie by phone	. . .	. . .	2
Gone to the movies	9	15	12
Watched a movie on a pay-cable channel (HBO, Showtime, etc.)	. . .	. . .	25
Watched a movie on a basic cable channel (A&E, TNT, TBS, USA, etc.)	. . .	. . .	38

	September 1990 (%)	*March 1991 (%)*	*September 1991 (%)*
Watched a movie on regular broadcast TV (ABC, CBS, NBC, PBS, Fox, independents, etc.)	. . .	. . .	63
Took snapshots	. . .	14	. . .
Had picture taken by someone	. . .	10	. . .
Used a video camera	. . .	5	. . .
N	(1,988)	(1,972)	(1,973)

During the past week, which of these things, if any, have you done for entertainment at home? [Card shown respondent.]

		Roper	
	February 1982 (%)	*February 1984 (%)*	*February 1990 (%)*
Watched a movie on TV	75	78	78
Watched some other kind of show on TV	70	67	73
Played a video tape or disc	6	7	29
Played an electronic game on your TV set	8	9	9
Read a book	. . .	46	44
Read a magazine	. . .	. . .	56
N	(2,000)	(2,000)	(1,977)

Now here is a list of a number of different things. [Card shown respondent.] Would you read down that list and call off each one you personally have done in the last 24 hours?

		Roper	
	August 1988 (%)	*August 1990 (%)*	*August 1991 (%)*
Used a computer at home	7	. . .	6
Used a computer at work	. . .	. . .	13
Used a facsimile machine[a]	. . .	7	8
Read a newspaper	75	65	. . .
Listened to news on radio	67	. . .	. . .
Watched news on TV	84	74	. . .
Listened to music on radio	69	. . .	. . .
Watched an entertainment show on TV	66	. . .	. . .
N	(1,987)	(1,997)	(1,983)

[a] In 1990, "used a facsimile machine at home or at work."

Videocassette Recorders

The questions in this section and the following four sections deal with the way that some of the new media technologies are used, and thus were, in most cases, asked only of respondents who had previously indicated that they owned or subscribed to a particular product or service. When it is reported, for example, that in July of 1980, 22 percent of respondents recorded less than 1 hour of TV on the VCR, this does not mean 22 percent of all respondents, but 22 percent of those who own VCRs (in this case, only 2 percent of the total sample). To emphasize this, I have, in each case, provided the base percentage on which the reported figures are calculated. The *N*s reported are my own calculation, derived by multiplying the base percentage by the total sample size, and are thus only approximate.

[Asked only of those who have a VCR:] About how many hours of TV shows per week would you say you usually record on your video cassette recorder?

Roper

	July 1980[a] *(%)*	*October–November 1983 (%)*	*October–November 1984 (%)*	*February 1985 (%)*
Less than 1 hour	22	31	24	27
1 to less than 2 hours	19	13	22	19
2 to less than 3 hours	11	12	15	16
3 to less than 4 hours	5	12	10	11
4 to less than 5 hours	3	2	8	5
5 to less than 6 hours	3	9	7	7
6 hours or more	14	15	8	11
Don't know/no answer	24	7	6	4
Base %	2	6	16	19
N	(40)	(60)	(320)	(380)

	September 1985 (%)	*September 1987 (%)*	*September 1989 (%)*
Less than 1 hour	20	40	36
1 to less than 2 hours	17	21	21
2 to less than 3 hours	20	12	14
3 to less than 4 hours	9	7	6
4 to less than 5 hours	11	4	4
5 to less than 6 hours	4	3	3
6 hours or more	12	6	7
Don't know/no answer	7	7	8
Base %	23	48	65
N	(459)	(955)	(1,282)

[a] In 1980, question read, ". . . you usually record on your video tape recorder?"

[Asked only of those who have a VCR:] When you use your VCR, do you usually eliminate the commercials, either by not recording them in the first place or by fast forwarding through them when you play the tape?"

		Roper	
	October–November 1984[a] *(%)*	*September 1985 (%)*	*September 1987 (%)*
Eliminate commercials	48	46	37
Let commericals run through	32	30	32
It varies[b]	14	15	17
Don't know	6	9	14
Base %	16	23	48
N	(320)	(459)	(955)

[a] In 1984, question read, "When you use your VCR, do you eliminate the commercials, or do they run through just as the on-air show did?"

[b] Volunteered.

[Asked only of those who have a VCR:] Do you ever rent movies to watch on your video cassette recorder? [If yes:] Looking at this card, which one category best expresses the number of movies you rent to watch on your VCR? [Card shown respondent.]

		Roper	
	September 1985 (%)	*September 1987 (%)*	*September 1989 (%)*
Yes, rent movies for VCR	87	91	86
Rent 4 or more per week	8	7	4
2 or 3 per week	16	16	16
About 1 per week	17	17	18
About 2 or 3 per month	28	28	24
1 per month	12	12	13
Less than one per month	6	11	10
No, do not rent movies	12	8	14
Base %	23	48	65
N	(459)	(955)	(1,282)

[Asked only of those who have a VCR:] Here is a list of different types of shows on television. [Card shown respondent.] Would you please read down that list and tell me which of these types of shows you tend to record most frequently on your video cassette recorder?

		Roper	
	September 1985 (%)	*September 1987 (%)*	*September 1989 (%)*
Full length movies	70	54	59
Sports events	39	24	26
Mini series, specials	33	20	25
Talk shows	6	6	9
Documentaries	22	19	19
Situation comedies	15	11	15
Daytime serials, soap operas	17	11	14
Weekly nighttime serials	14	9	10
Other nighttime dramas (police, detective, etc.)	17	10	11
Other kinds of shows or programs	21	22	20
None	7	17	20
Base %	23	48	65
N	(459)	(955)	(1,282)

[Asked only of those who have a VCR:] Here is a list of ways in which a VCR can be used. [Card shown respondent.] Would you please read down that list and call off all the ways in which the VCR in your household is used? Would you please read down that list again and tell me which of those things the VCR is used for most?

	Roper			
	VCR used		*Reason VCR used most*	
	September 1986 (%)	*September 1991 (%)*	*September 1986 (%)*	*September 1991 (%)*
To tape shows that are on while you are not home or too busy to watch	69	58	26	19
To tape TV shows that are on when there is something else you are watching on TV	56	40	11	6
To tape TV shows as you watch them	52	33	7	5
To tape TV shows that are on very late at night	45	28	5	4
To watch rented movies	88	88	54	60
To watch movies you bought or were given	54	51	6	9
To watch home movies made with a video camera	29	34	3	6
Base %	43	73	43	73
N	(859)	(1,440)	(859)	(1,440)

Cable TV

[Asked only of those subscribing to cable TV:] Which do you regard as the greatest advantage of cable—better reception and picture, or the special programs offered on cable TV, or the regular channels you wouldn't get without cable?

	Roper	
	January 1982 (%)	*July 1988 (%)*
Better reception and picture	20	17
Special programs	23	23
Additional regular channels	34	37
All three equal[a]	20	24
Don't know	6	2
Base %	35	56
N	(700)	(1,102)

[a] Volunteered.

[Asked only of those subscribing to cable TV:] Thinking of *all* the television you watched yesterday—not just cable—about how many hours did you personally spend watching any kind of TV? Now, thinking just of the programs you watched on cable television, like the ones on this card [card shown respondent], about how many hours did you personally spend watching programs on special cable channels?

	Roper (April 1982)	
	Any Kind of TV (%)	*Cable TV (%)*
None	12	34
Less than 1 hour	11	11
1 but less than 2 hours	20	18
2 but less than 3 hours	19	15
3 but less than 4 hours	14	9
4 but less than 5 hours	9	6
5 but less than 6 hours	6	2
6 but less than 7 hours	2	1
7 hours or more	5	2
Don't know	0	3
Base %	35	35
N	(700)	(700)

Computers

[All respondents:] Do you personally know how to operate a computer, or would you be interested in learning how to operate a computer, or is operating a computer something you want no part of?

			Roper		
	Already Know How (%)	*Want to Learn (%)*	*Want No Part Of (%)*	*Don't Know (%)*	*N*
January 1983	22	38	36	4	2,000
July 1984	29	38	29	4	2,000
July 1989	36	27	31	6	1,979

[Asked only of those who own a home computer:] Is your computer used for video games as well as for other purposes such as record keeping, problem solving, and word processing, or for games only, or for other purposes only?

		Roper	
	January 1983 (%)	*July 1984 (%)*	*July 1989 (%)*
Video games and other purposes	65	67	49
Games only	1	5	5
Other purposes only	29	26	42
Don't know	6	2	4
Base %	4	10	15
N	(80)	(200)	(297)

[Asked only of those who own a home computer:] Is this computer used for personal purposes, for business purposes, or both?

	Roper	
	July 1984 (%)	*July 1989 (%)*
Personal	59	35
Business	5	8
Both	29	48
Don't know	2	3
Computer used for games only	5	5
Base %	10	15
N	(200)	(297)

[Asked only of those who own a home computer:] Who in the household uses the computer for calculations and solving problems—both adults and children, or adults only, or children only?

	Roper		
	January 1983 (%)	*July 1984 (%)*	*July 1989 (%)*
Both adults and chldren	39	39	34
Adults only	49	42	49
Children only	4	13	8
Computer used for games only or don't know	7	9	9
Base %	4	10	15
N	(80)	(200)	(297)

Remote Control

[Asked only of those who have a remote control device for TV:] When you are watching TV and commercials come on, how frequently would you say you use the remote control to turn down the sound during the commercial? Would you say you turn down the sound during the commercials frequently, occasionally, rarely, or never?

	Roper (Base = 44%; N = 876)			
	Frequently (%)	*Occasionally (%)*	*Rarely (%)*	*Never (%)*
September 1987	35	21	16	28

[Asked only of those who have a remote control device for TV:] In which of these ways, if any, do you tend to use the remote control fairly often? [Card shown respondent.]

	Roper July 1988 (%)
Flip around the dial at program change time to see what is on	48
Flip to other channels during commercials	37
Flip to other channels from time to time during a program you are watching	19
Turn down the sound during commercials	18
None	7
Don't know	3
Base %	57
N	(1,122)

[Asked only of those who have a remote control device for TV:] How often, if ever, do you use the remote to "zip and zap" back and forth between programs while you are watching? Would you say all the time, a good deal of the time, hardly ever, or never?

	AIPO (Base = 79%; N = 979)			
	All the Time (%)	*A Good Deal (%)*	*Hardly Ever (%)*	*Never (%)*
August 1990	31	33	29	7

[Asked only of those who have a remote control device for TV:] How often do you use the remote to "zap" out ads, switch channels during ads, or turn the volume off during ads? Would you say all the time, a good deal of the time, hardly ever, or never?

	AIPO (Base = 79%; N = 979)			
	All the Time (%)	*A Good Deal (%)*	*Hardly Ever (%)*	*Never (%)*
August 1990	24	32	30	14

Answering Machines

[Asked only of those who own a telephone answering machine:] How often do you use your telephone answering machine to listen to who is calling and then decide whether to talk to them or call them back—do you do this frequently, often, occasionally, or never?

	Roper	
	December 1987 (%)	*December 1992 (%)*
Frequently	39	31
Often	14	17
Occasionally	27	27
Never	20	24
Base %	17	45
N	(337)	(890)

[Asked only of those who own a telephone answering machine:] Do you think a telephone answering machine actually saves you time?

	Black (Base = 24%; N = 193)	
	Yes (%)	*No (%)*
May 1987	57	40

A0610-1

Public Opinion Quarterly

Published for the American
Association for Public Opinion
Research by the University of
Chicago Press

MICHAEL L. ROBERTS, PEGGY A. HITE, and CASSIE F. BRADLEY — Attitudes toward Progressive Taxation

DAVID J. LANOUE and BARBARA HEADRICK — PM Popularity and Party Support in Britain

WILLIAM S. AQUILINO — Interview Mode Effects in Drug Surveys

GEORGE D. GASKELL, COLM A. O'MUIRCHEARTAIGH, and DANIEL B. WRIGHT — Response Alternative Effects and Vague Events

RESEARCH NOTES — Do Surveys Overstate Public Concerns? *by Arthur Sterngold, Rex H. Warland, and Robert O. Herrmann*

The Answering Machine Generation *by Robert W. Oldendick and Michael W. Link*

IN MEMORIAM — Don Cahalan, 1912–1992

THE POLLS — Review: Measuring Unemployment in the Nineties *by Janet L. Norwood and Judith M. Tanur*

Trends: Nuclear Power: Three Decades of Public Opinion *by Eugene A. Rosa and Riley E. Dunlap*

BOOKS — Reviews and Book Note *by Kurt Lang, John F. Dovidio, Ivor Crewe, and Maxwell McCombs*

Summer 1994

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR Advisory Committee

Howard Schuman, Chair (1993–97)
W. Russell Neuman (1990–94) Irving Crespi (1991–95)
J. Ronald Milavsky (1992–96)

Warren Mitofsky and Seymour Sudman, AAPOR Councillors-at-Large

Stanley Presser, *Editor*

Pia Peltola, *Office Manager*

The Polls
Reviews: Peter V. Miller, Albert E. Gollin
Trends: Robert Y. Shapiro

EDITORIAL BOARD

Herbert I. Abelson
Nathaniel Beck
Sandra H. Berry
Johnny Blair
Charles J. Brody
Pamela Conover
Myra Marx Ferree
Steven E. Finkel
Doris Graber
John P. Katosh
Gary King
Richard A. Kulka
Jacob Ludwig
William J. McGuire
Elizabeth Martin
Peter V. Miller
Diana C. Mutz
Lois Oksenberg
Vincent Price
Willard L. Rodgers
Nora Cate Schaeffer
Jacqueline Scott
Tom W. Smith
Charlotte Steeh
Seymour Sudman
John Sullivan
Marylee C. Taylor
Roger Tourangeau
Michael W. Traugott
Michael Weeks
Susan Welch
John Zaller

Public Opinion Quarterly (ISSN 0033-362X) is published quarterly in Spring, Summer, Autumn, and Winter, by The University of Chicago Press, 5720 S. Woodlawn Ave., Chicago, IL 60637. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $50.00; individuals 1 year $23.00. Special rates available to members of AAPOR. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $20.00. In Canada, add 7% GST to subscription price. Outside of U.S.A., add $4.00 for each year's subscription to cover postage. Japanese subscription agent: Kinokuniya Company Ltd. Individuals have the option to order directly from the University of Chicago Press. Single copy rates: institutions $12.50; individuals $5.75.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to Journals Fulfillment Department, University of Chicago Press, P.O. Box 37005, Chicago, IL 60637. FAX: (312) 753-0811.

Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Sociology Department,
University of Maryland, College Park, MD 20742-1315
Business, Subscriptions, Back Issues: University of Chicago Press,
P.O. Box 37005, Chicago, IL 60637
AAPOR: P.O. Box 1248, Ann Arbor, MI 48106

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

© 1994 by the American Association for Public Opinion Research

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ∞

PUBLIC OPINION QUARTERLY

VOLUME 58 / SUMMER 1994 / NUMBER 2

Articles

Understanding Attitudes toward Progressive Taxation 165
MICHAEL L. ROBERTS, PEGGY A. HITE, and CASSIE F. BRADLEY

Prime Ministers, Parties, and the Public: The Dynamics of Government Popularity in Great Britain 191
DAVID J. LANOUE and BARBARA HEADRICK

Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment 210
WILLIAM S. AQUILINO

Survey Questions about the Frequency of Vaguely Defined Events: The Effects of Response Alternatives 241
GEORGE D. GASKELL, COLM A. O'MUIRCHEARTAIGH, and DANIEL B. WRIGHT

Research Notes

Do Surveys Overstate Public Concerns? 255
ARTHUR STERNGOLD, REX H. WARLAND, and ROBERT O. HERRMANN

The Answering Machine Generation: Who Are They and What Problem Do They Pose for Survey Research? 264
ROBERT W. OLDENDICK and MICHAEL W. LINK

In Memoriam

Obituary for Don Cahalan, 1912–1992 274
JACK ELINSON, HELEN M. CROSSLEY, ANNE SCHUETZ ZANES, ANN F. BRUNSWICK, SELMA F. MONSKY, and DEAN I. MANHEIMER

The Polls

Review: Measuring Unemployment in the Nineties 277
JANET L. NORWOOD and JUDITH M. TANUR

Trends: Nuclear Power: Three Decades of Public Opinion 295
EUGENE A. ROSA and RILEY E. DUNLAP

Book Reviews

W. Russell Neuman, Marion R. Just, and Ann N. Crigler, *Common Knowledge: News and the Construction of Political Meaning* 325
KURT LANG

Paul M. Sniderman, Richard A. Brody, and Philip E. Tetlock, *Reasoning and Choice: Explorations in Political Psychology* 327
JOHN F. DOVIDIO

Helmut Norpoth, *Confidence Regained: Economics, Mrs. Thatcher and the British Voter* 329
IVOR CREWE

Leo Bogart, *Preserving the Press: How Daily Newspapers Mobilized to Keep Their Readers* 331
MAXWELL MCCOMBS

Book Note: Shanto Iyengar and William J. McGuire, eds., *Explorations in Political Psychology* 333

Notice to Contributors

Prospective contributors should refer to the detailed Notice to Contributors that appears in the Spring 1994 issue of *Public Opinion Quarterly*.

Advertising space in *Public Opinion Quarterly* is available, as is rental of its subscriber list. For information and rates, please contact the advertising sales staff, University of Chicago Press, Journals Division, 5720 S. Woodlawn Ave., Chicago, IL 60637. Advertising and list rental are limited to material of scholarly interest to our subscribers.

THE ANSWERING MACHINE GENERATION

WHO ARE THEY AND WHAT PROBLEM DO THEY POSE FOR SURVEY RESEARCH?

ROBERT W. OLDENDICK
Institute of Public Affairs, University of South Carolina
MICHAEL W. LINK
Institute of Public Affairs, University of South Carolina

Abstract Increased incidence of telephone answering machines and the use of such devices to screen calls pose a potential threat to the representativeness of samples in telephone surveys. Using data from nine statewide surveys, this analysis examines the extent to which answering machines are used to screen calls and the demographic characteristics associated with answering machine use and call screening. Results show that at most two to three percent of households use answering machines to screen calls, and that such screening is more likely to take place in households with higher family incomes, outside rural areas, and which include individuals who are younger and have higher levels of education. While call screening does not presently threaten the representativeness of samples in telephone surveys, the increased incidence of answering machines together with the increased percentage of households indicating these devices are sometimes used to screen calls demonstrate that the potential bias from this source is growing.

Although there are many potential sources of error in surveys, among the most disturbing are those that threaten the representativeness of the sample. Lack of representativeness resulting from noncoverage, refusals to participate, or other forms of nonresponse jeopardize the accuracy of survey results. In this study we address one aspect of this problem by examining the extent to which the increased use of tele-

This research was supported, in part, by an award from the University of South Carolina's Research and Productive Scholarship Fund. The authors would like to thank the anonymous referees who supplied many valuable suggestions.

Public Opinion Quarterly Volume 58:264–273 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5802-0007$02.50

phone answering machines may be contributing to declining response rates.

Previous work in this area has focused primarily upon the incidence of answering machines (Baumgartner 1990; Tuckel and Feinberg 1991), strategies for contacting these households (Piazza 1993; Piekarski 1990), and demographic correlates associated with answering machine use (Oldendick 1993). Most of this research indicates that, although a significant percentage of households have such machines, a substantial proportion of these households are nonetheless accessible to survey researchers.

The current investigation extends this research by examining trends in answering machine use and by exploring more extensively the correlates of both answering machine ownership and the use of these machines to screen calls.

Data and Methods

The data come from nine random digit dialed telephone surveys of the adult (18 and over) population in South Carolina. Respondents within households were selected using the last birthday method (Oldendick et al. 1988; O'Rourke and Blair 1983). Households were contacted a minimum of six times before a final disposition was assigned. Senior interviewers recontacted households in which an interview was initially refused in an attempt to obtain a completed interview. After two refusals, no additional contacts were made. The surveys were conducted between November 1989 and November 1992 and covered issues such as public transportation, state budget priorities and taxes, race relations, the environment, and the arts.[1]

The prevalence of telephone answering machines was estimated in two ways. First, telephone numbers that produced any contact with an answering machine were tracked to determine whether they resulted in completed interviews, were never contacted, or were otherwise disposed. Second, the final questions in each of these surveys were (1) "Does your household have a telephone answering machine?" and, for those households that reported having one, (2) "Does anyone in your household ever use this telephone answering machine to screen unwanted calls?"

---

[1] Response rates for these surveys ranged from 60.2 percent for the September 1990 survey to 66.5 percent for the one conducted in March 1990. Response rates were calculated by dividing the total number of completions + partial completions by the number of completions + partial completions + refusals + ill/senile/not available during fielding period + the estimated number of households among the never answered numbers.

Prevalence of Answering Devices: The Potential Problem

Previous studies have estimated the incidence of answering machines in households nationally to be approximately 25 percent (Parsley 1992; Piekarski 1990; Tuckel and Feinberg 1991).[2] As Tuckel and Feinberg (1991, p. 205) explain, in estimating the accessibility of these households it is important to examine the difference between "the observed frequency of telephone answering machines as a response disposition and the frequency we would expect if all answering machine owners were to use their machines to the maximum extent." Following their lead, we begin with the final dispositions of the sampled numbers for each survey.

Table 1 presents the percentage of calls resulting in a final "answering machine" disposition—that is, each time these numbers were dialed they went unanswered or reached an answering machine. While this outcome shows a slight tendency to increase over time, the variation is not statistically significant. Overall, call attempts resulted in either consistent answering machine contacts or a combination of answering machines and "no answers" about 2.5 percent of the time. At most, then, about 2 percent of households may be using a telephone answering machine to screen calls on a consistent basis.[3]

Self-Reports of Answering Machine Ownership and Screening

While the use of answering machines to screen calls may not be a major problem at present, other data show that both answering machine ownership and screening—at least of a sporadic nature—are increasing. Table 2 shows that the incidence of self-reported ownership of these machines in South Carolina has increased from 25 percent in 1989 to approximately 40 percent in 1992. More important, there has

---

[2] A more recent estimate places this figure at more than 43 percent (Electronic Industries Association 1991).

[3] This figure is somewhat lower than the 5.7 percent reported by Tuckel and Feinberg (1991, pp. 206–7). While part of this difference may result from the fact that their study involved a different population, it is more likely a result of the use of alternative procedures. In their study, most households were called a maximum of three times, while in the surveys reported here numbers were attempted a minimum of six times. In addition, unlike our study, their "answering machine" disposition included not only consistent answering machine contacts, but also those cases in which a number yielded a callback followed by an answering machine on a subsequent call. Our estimate of 2 percent is an upper bound in that not all these households may be "screening" calls, in the sense that someone is at home, but using the answering device to determine who is calling; there are likely to be instances in which no one was available at the times we called, even when 20 or more calls were attempted.

Table 1. Sampled Telephone Numbers Resulting in Contact with an Answering Machine or Combination of Answering Machine and No Answer

	N[a]	% Total[b]	% Estimated Households[c]
November 1989	20	1.5	2.2
March 1990	26	1.4	2.1
September 1990	26	1.4	1.9
October 1990	51	1.1	1.6
April 1991	37	1.8	2.7
October 1991	39	2.2	2.9
November 1991	11	.9	1.2
April 1992	39	2.4	3.1
November 1992	47	2.8	3.6

[a] Number of all sampled phone numbers that resulted in an answering machine disposition on each attempt or a combination of answering machine and "no answer" dispositions.

[b] Percentage of all sampled numbers that resulted in an answering machine disposition on each attempt or a combination of answering machine and "no answer" dispositions.

[c] In calculating the "% estimated households," the proportion of "never answered" numbers estimated to be households (rather than "ring no answer" numbers) was assumed to be the same as that for those numbers for which a household determination was made.

also been an increase—from 38 percent to 48 percent—in the percentage of people indicating that someone in the household sometimes uses the device to screen calls. Since total call screening is a product of these two variables, it rose during this period from 9 percent to 19 percent. While these numbers are relatively large in comparison to the small percentage of households where consistent answering machine use was found, they are tempered by the broad scope of the question, "Does *anyone* in your household *ever* use this telephone answering machine to screen unwanted calls?" Obviously such screening was not occurring all the time, or we would never have reached them.[4]

The finding of little change in the percentage of households in which a telephone answering machine is consistently encountered seems somewhat in conflict with the report of significant increases in the proportion of households with answering devices and in the percentage of such households that use these devices to screen unwanted calls.

---

4. When interviewers in these studies reached an answering machine, they terminated the call without leaving a message.

Table 2. Trends in Households with Answering Machines—State Data

	November 1989	March 1990	September 1990	October 1990	April 1991	October 1991	November 1991	April 1992	November 1992
% with answering machines	25 (138)	26 (205)	28 (217)	27 (539)	34 (280)	37 (295)	35 (206)	40 (313)	39 (314)
% of owners who use for screening	38 (53)	42 (85)	44 (96)	41 (222)	55 (155)	48 (142)	46 (93)	47 (146)	48 (152)
% of total sample who use for screening	9 (53)	11 (85)	12 (96)	11 (222)	19 (155)	18 (142)	16 (93)	19 (146)	19 (152)

NOTE.—Percentages are based on households in which an interview was completed. Numbers in parentheses are *N*'s.

While these concepts are related, they are distinguished in that the former reports on recent telephone survey experience in encountering the answering machine barrier while the latter estimates the size of the potential barrier.

Demographic Correlates of Ownership and Screening

Pooling across the nine surveys, table 3 shows the relationship of six background variables to answering machine ownership and to use of the machine for call screening.[5] In examining these data it should be kept in mind that our questions measure incidence and use at the household level. It is not appropriate, therefore, to infer that the individual-level characteristics described here are directly related to answering machine ownership or the use of such devices to screen calls. Rather, this analysis indicates the individual characteristics of those living in households that (1) have an answering machine and (2) use it for screening.

The data indicate that family income is significantly correlated with ownership. Ownership is much more prevalent among those with higher incomes, and the percentage of households with answering machines increases steadily across income groups.

Urbanicity also was significantly related to ownership. Suburbanites were most likely to report having an answering machine, followed by those living in urban areas.

Turning to characteristics of individuals living in households with a telephone answering machine, education, age, and race each show significant relations with ownership. Higher educated individuals, particularly those with at least some college, more often reported an answering machine in their household. Similarly, both younger respondents and whites were more likely to report an answering machine.

While these differences in answering machine ownership across subgroups are important, a more critical consideration from a survey research perspective is the extent to which such devices are used to screen calls. With the exception of age, however, these variables are

---

[5] The data are pooled since the relations between these characteristics and answering machine ownership did not change over time. Given the large number of cases in this pooled data set, all the reported relationships were statistically significant at the .0001 level. More important, the relationship between answering machine ownership and total family income, urbanicity, education, and age was significant at the .001 level in each of the nine surveys. The relationship between ownership and race was significant at the .001 level in seven of the nine surveys; it was significant at the .01 level in November 1991 and not significant ($p > .05$) in September 1990. The relationship between answering machine ownership and household size was not significant at the .01 level in any of the surveys. Data for the individual surveys are available from us on request.

Table 3. Demographic Correlates of Telephone Answering Machine Ownership and Use for Screening Calls

	% with Answering Machine	% Who Use to Screen Calls	*N*
Total sample	31	14	7,649
Household characteristics:			
Total family income:			
Under $15,000	15	6	1,600
$15,000–29,999	30	14	2,021
$30,000–49,999	40	19	1,680
$50,000 and over	54	25	1,197
Household size:			
One	29	15	1,157
Two	34	16	1,919
Three	38	19	1,168
Four or more	32	14	1,647
Urban/rural area:			
Urban	33	16	1,476
Suburban	43	20	1,696
Rural	26	12	2,031
Individual characteristics:			
Education:			
Less than high school diploma	16	5	1,772
High school diploma	25	11	2,600
Some college	38	18	2,005
College degree	49	23	1,540
Age:			
18–29	36	18	2,389
30–44	38	19	2,206
45–64	30	12	1,966
65 and over	11	2	1,049
Race:			
Blacks	19	9	2,182
Whites	36	16	5,661

not significantly related to screening calls among those households with answering machines (data not shown).[6] Thus to the extent that call screening is related to background characteristics (see table 3), it is mainly due to the higher incidence of answering machine ownership among certain groups rather than to differences between groups in likelihood of screening.[7]

Returning to the question of the subgroups that are most likely to be missed due to the use of answering machines as screening devices, table 4 shows the joint effect of education and age on answering machine ownership and screening.[8] Cell entries are the deviations from the overall percentage of the sample using an answering machine to screen calls (14 percent). A positive figure means the group is more likely to screen calls, while a negative entry means a group is less likely to screen.

Within each age group, those with some college or a college degree were more likely than those with less education to screen calls. Moreover, within each of these education groups, younger respondents were generally more likely than older respondents to screen calls. Thus answering machine screening is a larger potential obstacle among certain subgroups, particularly younger, more educated respondents. Piazza (1993) reports a similar finding with respect to education.

Conclusions

This analysis provides a mixed message for survey research. The 2–3 percent of households that may regularly use answering machines to screen calls do not now represent a serious threat to the representativeness of telephone survey samples. On the other hand, both answering machine ownership and the use of these devices for at least some call screening increased consistently over the 3-year period of this study. The potential threat to representativeness is, therefore, also increasing.

Incidence of answering machine use and screening is also greater among certain subgroups of the population. Those with higher family incomes and more education, younger people, whites, and those living

---

[6] Significant differences across age groups were noted in three of the nine surveys, with these differences largely a result of a much lower incidence of screening among those 65 and older.

[7] The relationships between these characteristics and use of answering machines for screening did not change over time.

[8] Education and age were chosen on the basis of a multivariate logit analysis, which identified them as the most important factors affecting answering machine ownership and call screening. This analysis is available from us on request.

Table 4. Age and Education Subgroups Most Likely to Screen for Unwanted Calls

	Age				
Education	18–29	30–39	40–49	50–59	60+
Less than high school	−6	−5	−8	−5	−12
High school degree	0	0	−2	−4	−10
Some college	+7	+8	+6	+1	−9
College degree	+17	+12	+10	+8	−8

NOTE.—Figures represent the percentage above or below the total sample percentage of 14 percent. The smallest number of cases in any cell is 171 (those age 40–49 with less than a high school education).

in nonrural areas are more likely to own a telephone answering machine and to screen calls. (We should note that the population on which this analysis is based does not include any very large metropolitan areas, where use of telephone answering devices for screening may be higher.)

Strategies suggested by previous studies, such as repeated calls or leaving messages on answering machines, may become less effective as the incidence of these machines and their use for screening calls becomes more widespread. Thus survey researchers need to track not only the incidence of answering devices and the demographic factors associated with screening, but also the effects that strategies such as leaving various types of messages have on reaching the answering machine generation.

References

Baumgartner, Robert M. 1990. "Telephone Answering Machine Messages and Completion Rates for Telephone Surveys." Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA.

Electronic Industries Association. 1991. *Consumer Electronics U.S. Sales*. Electronic Industries Association, Washington, DC.

Oldendick, Robert W. 1993. "The Effect of Answering Machines on the Representativeness of Samples in Telephone Surveys." *Journal of Official Statistics* 9:663–72.

Oldendick, Robert W., George F. Bishop, Susan B. Sorenson, and Alfred J. Tuchfarber. 1988. "A Comparison of the Kish and Last Birthday Methods of Respondent Selection in Telephone Surveys." *Journal of Official Statistics* 4:307–18.

O'Rourke, Diane, and Johnny Blair. 1983. "Improving Random Respondent Selection in Telephone Surveys." *Journal of Marketing Research* 20:428–32.

Parsley, Teresa. 1992. "Prevalence and Use of Answering Machines: Data from a Nationwide Telephone Survey." Paper presented at the Field Directors and Field Technologies Conference, St. Petersburg, FL.

Piazza, Thomas. 1993. "Meeting the Challenge of Answering Machines." *Public Opinion Quarterly* 57:219–31.

Piekarski, Linda. 1990. "Answering Machine Households Not So Elusive." *The Frame* 13:1.

Tuckel, Peter S., and Barry M. Feinberg. 1991. "The Answering Machine Poses Many Questions for Telephone Survey Researchers." *Public Opinion Quarterly* 55:200–217.

THE POLLS—A REVIEW*

MEASURING UNEMPLOYMENT IN THE NINETIES[1]

JANET L. NORWOOD
JUDITH M. TANUR

At the beginning of 1994, the measured national unemployment rate jumped upward because the survey questionnaire used to measure unemployment had been redesigned and data collection methods changed. In November 1993, government information suggested that the jobless rate, undoubtedly one of the most politically sensitive of all federal statistics, had been substantially underestimated—by about half a percentage point—for many years. And this, in spite of the fact that the Current Population Survey (CPS), the 53-year-old survey that produces the data to assess the unemployment rate and provides a great deal of other labor force information, has long been used as a model of the best in survey design both in this country and around the world. What had changed?

The CPS had undergone the largest redesign in the survey's history. The redesign program, begun in 1986, used cognitive techniques to revamp the questionnaire to reflect labor market behavior better than it had in the past and to incorporate advances in data collection. Although it had been customary to redesign the CPS sample and change

JANET L. NORWOOD is senior fellow at the Urban Institute in Washington, DC, and served as the U.S. commissioner of labor statistics from 1979 through 1991 and consequently had responsibility for the first several years of the CPS redesign project. JUDITH M. TANUR is professor of sociology at the State University of New York at Stony Brook.

**This review differs from those appearing previously in Poll Reviews, which was established in 1986 as a forum (analogous to book reviews) for the critical examination of published survey research and public polls. This article, originally solicited by us and subjected to standard editorial review, was written by two social scientists who were intimately involved with the processes they describe. It offers an authoritative "inside" appraisal of both context and rationale for the revision of an important government survey, the Current Population Survey.—The Poll Review Editors*

[1] This article was written in December 1993. It is based on the estimates of the differences between the overlap sample for the redesigned CPS and the unrevised CPS that were available at that time. Researchers for BLS and Census have prepared a great many papers that describe the redesign we discuss in this review. A "CPS Redesign Bibliography" is available from Paul Mullin at BLS. Our discussion draws on this bibliography as well as BLS briefing materials.

Public Opinion Quarterly Volume 58:277–294 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5802-0006$02.50

the population estimates used in poststratification after completion of each decennial census, the CPS questionnaire had remained essentially the same since 1967.

The options for the redesign for the 1990s were quite clear. The agencies could focus on design of a new sample, updating population estimates, and computerizing data collection, or they could have a more comprehensive redesign and modernize the questionnaire as well. Since the CPS is redesigned only once a decade, unless questionnaire issues were addressed this time, their change probably could not be implemented until the next redesign after the 2000 census. Although work with the questionnaire would raise the redesign budget, the benefits from questionnaire improvement were judged to be worth the incremental cost.

So, the goals for the redesign were established: to clarify the labor force concepts used in the survey, to make a few small changes in definition, to reduce the potential for response error, to expand the data available from the survey, to minimize the burden on respondents, and to put into effect a fully computerized survey. Decisions on particular improvements used in the redesign were based on the experience gained in a series of research projects undertaken in a joint effort by the U.S. Bureau of Labor Statistics (BLS), the survey sponsor, and the U.S. Bureau of the Census (Census), which conducts the survey. Early in the redesign planning, the two agencies decided to provide for an overlap sample so that comparisons could be made between the data generated by the old survey and those by the new one. Thus, to provide this estimate of the effect of the changes on the data collected, a separate national sample survey using the new questionnaire and technology was begun in July 1992 and carried through December 1993.

The Issues

The experience accumulated by BLS in analyzing employment and unemployment data from the CPS pointed to several ways in which the current questionnaire could be improved. Several of these areas for potential improvement arose from changes in society and in survey research methods. First, the labor market had changed radically in the last 30 years or so. The nation's industry had been restructured; women, especially mothers of small children, had become major participants in the labor force; the proportion of minorities had increased markedly; and the number of part-time and temporary workers had risen considerably. Some of the questions needed to be revised to ensure that these changes were properly reflected in the survey, and some questions, especially those dealing with job search, workers on layoff expecting to be recalled, and discouraged workers, needed special attention. Second, the country had more older people, some of whom worked and some of whom were retired. Alternative approaches to surveying this part of the population should be incorporated into the survey. Third, new advances in methods of questionnaire design could tailor questions more closely to respondents' and to survey designers' concepts, and changes in data collection technology permitted the development of a fully automated survey with the flexibility to use alternative questionnaire approaches not possible with a paper-and-pencil operation.

Another issue involved the definition of unemployment and the capacity to operationalize it in the survey. Since its inception more than half a century ago, the CPS has divided the population into three groups: those who work (the employed), those without jobs who are looking for work or laid off from jobs but expecting to be recalled (the unemployed), and those who are outside the labor force entirely. A person is considered employed if he or she worked for pay or profit during the survey week—even if only for 1 hour—or put in 15 hours or more of unpaid work in a family farm or business or had a job but was not at work. The unemployed include all those who (*a*) were not employed during the survey week; (*b*) were available for work and could have taken a job if one were offered; and (*c*) have engaged in some active job search during the preceding 4 weeks unless laid off with expectation of recall. This "activity-based" definition for unemployment was reviewed by two presidential commissions and has been adopted by most other developed countries of the world.[2]

But experience had turned up problems in several questions with the CPS that are used to classify people into the three labor force categories—the meaning of layoff, the manner in which respondents thought about and reported unpaid family work and part-time work, and what respondents understood job search to be. (Questions about discouraged workers, and those on income and hours of work, although not directly applicable to classification into labor force categories under the official definition, were also thought to need review.)

The layoff question, for example, directly affects the number counted as unemployed since a worker on layoff is assumed to be waiting for recall and hence is counted as unemployed without any requirement for job search. (The different practices of countries in the classification of persons "laid off" in the unemployment statistics of the Organization for Economic Cooperation and Development [OECD] countries had been discussed in meetings as early as 1979, when practices in determining workers' attachment to individual employers were examined in European countries and in the United States and Canada [see Grais 1979; U.S. Bureau of Labor Statistics and U.S. Census Bureau 1986].) Thus, the CPS's official meaning of "on layoff" requires an expectation of recall to the job within a short time, but there was some suspicion that the term was used by respondents as a euphemism for "fired" (see Palmisano 1989).

---

[2] The first of these presidential commissions, appointed by President Kennedy, was chaired by Robert Gordon and reported in 1962 (see President's Committee to Appraise Employment and Unemployment Statistics 1962). The second commission, appointed by President Carter, was chaired by Sar Levitan and reported in 1979 (see National Commission on Employment and Unemployment Statistics 1979).

Bureau of Labor Statistics and Census Bureau analysts also believed that the operational definition of looking for work in the existing CPS questionnaire needed clarification and that too much discretion on what constitutes job search was left to the interviewer and the respondent. They were also concerned that the very first CPS labor force question, "What was . . . doing most of LAST WEEK—working, keeping house, going to school or something else?" was often not read to respondents in full, thus possibly influencing their choices (see U.S. Bureau of Labor Statistics and U.S. Census Bureau 1986).

Yet another issue involved the recommendations of the 1979 Levitan Commission. Although the BLS had already incorporated several commission recommendations that did not require costly testing (e.g., inclusion of the armed forces in the labor force in 1983 and collection of monthly data on school enrollment, first published in 1985), funds had not been available for the research and testing required for implementation of two other important changes. The commission had recommended (1) that respondents be asked if they usually worked more hours than they did in the survey week, and, if so, the reason for the shorter hours in that week and (2) making the definition of discouraged workers less dependent on the worker's opinion by requiring prior job search. Neither of these recommendations had been incorporated although two secretaries of labor, Ray Marshall and Ray Donovan, had approved them.

Finally, both agencies believed that the survey could be considerably improved through the use of cognitive principles in research on alternative approaches to dealing with these issues. The Census Bureau had previously had experience with laboratory experiments in testing survey forms, and BLS had recently established a Collection Procedures Research Laboratory (Dippo and Norwood 1992). Both agencies were also interested in applying modern computerized techniques to survey data collection, processing, and analysis.

The Process

The leadership of the two agencies began a joint effort to develop a comprehensive strategic plan to reaffirm the CPS's leadership in labor force surveys by using state-of-the-art questionnaire design and advanced technology. A series of meetings with the senior staff of the two agencies, including the BLS commissioner and the Census director, delineated the comprehensive set of tasks that needed to be carried out. Several joint task forces were established to develop the research issues that needed to be addressed and to establish the necessary plans, budget estimates, and time tables for implementation.

The work began with efforts to identify areas of the old questionnaire that might need change. First, experts from BLS and the Census, as members of the CPS Questionnaire Redesign Task Force, reviewed problem areas. Based on their experience in developing and analyzing the data, they were often able to see questions that had a high potential for response error or that failed to implement existing definitions properly. They also noted questions needing definitional changes and identified areas of labor force analysis for which more data were needed. They believed that because of the dramatic increase in labor force participation of women and the rise in part-time and temporary workers, more information was needed from the survey about these categories of workers. Bureau of Labor Statistics staff were also eager to develop a more user-friendly processing system so that these and other important analytical issues could be addressed more easily and more quickly than in the past.

Next, the task force recommended a program of preliminary research. Respondents were brought together in focus groups to explore their understanding of the concepts underlying the questions asked. It was this technique, for example, that confirmed a discrepancy between the respondents' and the CPS's use of the term "on layoff." Another preliminary technique tapped interviewers' experience with other CPS concepts by using card sorts to explore their coding of open-ended questions about job search and other matters (Fracasso 1989).

The next step was to design two new questionnaires in order to evaluate different proposed solutions for the problems discovered in the preliminary analyses. With the use of a Computer Assisted Telephone Sample (CATI) developed through Random Digit Dialing (RDD) of 72,000 households, these two alternatives were tested against the standard CPS interview, with households randomly assigned to one of the three questionnaires.

Several evaluation methods were used during and after this test in an attempt to develop the single best alternative questionnaire for testing against the standard one in the final phase of development. Again the interviewers' expertise and experience provided data, as 68 interviewers answered a self-administered questionnaire dealing with their views on such issues as (*a*) the most difficult questions for interviewers to ask or for respondents to answer, (*b*) misunderstood concepts, and (*c*) questions with high refusal rates. Interviewers also met in focus groups of eight to 10 people to discuss these and other sources of difficulty in administering the three questionnaires.

Next, in respondent debriefings, follow-up questions were asked to determine whether terms or phrases were misunderstood, whether respondents' understanding of key CPS concepts matched official definitions, and whether certain questions were superfluous. In addition, vignettes, brief scenarios depicting activities that might or might not be consistent with a CPS concept (e.g., looking for work) were presented to check respondents' understanding of these concepts. Both the follow-up probes and the vignettes were keyed to the respondents' answers to the main questionnaire in the sense that, for example, only those reporting a household member as looking for work would be asked about the looking-for-work concept.

In direct observation of interviews, Census and BLS researchers monitored 229 interviews while they were in progress and coded interviewer and respondent behaviors. This work used forms and guidelines adapted from the work of Cannell et al. (1989).

Finally, item nonresponse rates were examined to see if they differed for comparable questions across forms of the questionnaire and response distributions were compared to see if comparable question sets yielded different results because of different wording or sequencing.

Several papers have detailed this extensive pretesting (e.g., Esposito et al. 1992; Polivka and Rothgeb 1993a, 1993b). Essentially, the two versions of each question were examined using the four evaluation techniques detailed above. The researchers weighed the evidence from all of the techniques to choose the preferred version of each question. These were combined into an almost final version of the questionnaire that was further field tested against the standard questionnaire for fine tuning.

Changes in the Questionnaire

Perhaps the most profound change in the questionnaire, and the one that is probably most responsible for the increase in the measured unemployment rate, is in the very first set of questions that ask about work in the reference week. They are crucial for the classification into the three categories of employed, unemployed, and out of the labor force. The old and new versions of these questions appear in table 1. Experts at the two agencies believed that this first group of survey questions in the old CPS might be misunderstood by respondents, that volunteer activities might be improperly identified (the criterion is work for pay), and that temporary work for pay that should be included might be omitted. The old version also seemed to encourage interviewer discretion in the order of presentation of the alternatives in the first question and required that the respondent volunteer information about a family business or farm in order for the next question to include an inquiry about unpaid work. The new questionnaire first establishes the existence of a family business or farm, tailors the question to that reality, and makes the first stab at classification into the labor force categories by asking about any work for pay (cf. numbers 1–4 under Revised Questionnaire with numbers 1 and 2 in the Unrevised CPS col. in table 1).

Next, the new questionnaire uses cognitive principles to make the respondent's task easier and to encourage the reporting of more accurate information. For example, in the old questionnaire, the respondent is first asked the actual hours he or she worked last week and only then (in some cases) the usual hours worked. This seems to be an effort to avoid "influencing" respondents or encouraging them to give stereotypical answers to questions about the reference week. In the new questionnaire, the respondent is asked first how many hours he

Table 1. Employment and Unemployment Questions

Revised Questionnaire	Unrevised Current Population Survey
1. Does anyone in this household have a business or a farm?	
2. *Last Week* did you do *any* work for (either) pay (or profit)? Parenthetical filled in if there is a business or farm in the household. If 1 is "yes" and 2 is "no," ask 3.	1. What were you doing most of *Last Week?* (working or something else?) (keeping house or something else?) (going to school or something else?)
3. *Last Week*, did you do any unpaid work in the family business or farm? If 2 and 3 are both "no," ask 4.	If answer indicates "with a job, but not at work" (either temporarily or on layoff), ask 2, and if 2 is "no" ask 4. If answer indicates "working," skip 2. All others, ask 2.

Table 1 (*Continued*)

Revised Questionnaire	Unrevised Current Population Study
4. *Last Week*, (in addition to the business,) did you have a job, either full- or part-time? Include any job from which you were temporarily absent.	2. Did you do any work at all *Last Week*, not counting work around the house? (Note: If farm or business operator in household, ask about unpaid work.)
Parenthetical filled in if there is a business or farm in the household.	3. Did you have a job or business from which you were temporarily absent or on layoff *Last Week*?
If 4 is "no," ask 5.	If "no," ask 5. If "yes," ask 4.
5. *Last Week*, were you on layoff from a job?	4. Why were you absent from work *Last Week*?
If 5 is "yes," ask 6. If 5 is "no," ask 8.	
6. Has your employer given you a date to return to work?	
If "no," ask 7.	
7. Have you been given any indication that you will be recalled to work within the next 6 months?	
If "no," ask 8.	
8. Have you been doing anything to find work during the last 4 weeks?	5. Have you been looking for work during the past 4 weeks?
If "yes," ask 9.	If "yes," ask 6.
9. What are all of the things you have done to find work during the last 4 weeks?	6. What have you been doing in the last 4 weeks to find work?
Individuals are classified as employed if they say "yes" to questions 2, 3 (and work 15 hours or more in the reference week or receive profits from the business/farm), or 4. Individuals who are available to work are classified as unemployed if they say "yes" to 5 and either 6 or 7, or if they say "yes" to 8 and provide a job search method that could have brought them into contact with a potential employer in 9.	Individuals can be classified as "at work" at question 1 or 2. Individuals can be classified "employed, temporarily absent" with the combination of 1 and 4, or 3 and 4. Individuals who are available to work can be classified as unemployed with combinations of 1 and 4, 3 and 4, 1 and 6, or 5 and 6.

SOURCE.—U.S. Bureau of Labor Statistics 1993.

or she usually works and then is encouraged to adjust this initial estimate for the extra hours worked or for time off in the reference week, thus invoking the principle of anchoring and adjustment. Care was taken with the new questionnaire to keep the questions neutral in order to avoid influencing the response, but questions that seek factual information can be tailored in this way to offer the respondent a useful estimation strategy. (See, e.g., Jabine et al. [1984], where the application of anchoring and adjustment and of other cognitive strategies for survey questionnaires is discussed.)

Another example is the way in which the new questionnaire guides people to think about the periodicity of their activities before answering questions affected by their timing. In questions about earnings, for example, the old questionnaire required most respondents to report weekly gross pay. In the new questionnaire, the respondent is first asked whether it would be easiest to report earnings on an hourly, weekly, annual, or some other basis. The respondent then reports for the time period that seems most natural, and the computer can do whatever adjustment is necessary.

Next, the "through the looking glass" stance—the view that the CPS words mean whatever the questionnaire writers say they mean—has been abandoned. If the CPS definition of "on layoff" implies the expectation of recall, contrary to the everyday usage of the term, the new questions asking about layoff now specifically query the expectation of recall.

Yet another change—this one facilitated by the use of the new technology—is dependent interviewing. The CPS had, for the most part, been conducted as though it were a new survey every month despite the fact that it is a rotating panel survey (households are in the sample for 4 months, out for 8 months, and then back in for 4 months before leaving the sample entirely). With the old (paper-and-pencil) questionnaire, each monthly CPS interview began anew, and the interviewer had no access to the responses to the labor force questions of the previous month. Although a survey control card is prepared at the time of the first interview, the control card does not contain data from earlier responses for particular questions. In fact, the interviewer had no direct way of knowing what the respondent had replied to each of the questions in the previous month. While this information is not necessary for all questions, the response given in the previous month could, in some cases, help the interviewer to verify the accuracy of the response or to help in coding it.

In the redesign, dependent interviewing is used sparingly for information that is especially hard to standardize. An example can be seen in tracking the month-to-month changes in occupation and industry. Such changes have been especially volatile and often spurious. This is because in the old questionnaire the subject's job was described anew each month and the problems of coding these open-ended responses often caused coders to come up with different categorizations across months even in the absence of real change. With the advent of computer-assisted interviewing (both computer-assisted telephone interviewing [CATI] and computer-assisted personal interviewing [CAPI]), the new CPS can easily use dependent interviewing: if the interviewer determines that the subject is still working for the same firm, he or she then asks if the subject is in the same job as he or she was in the previous month. An affirmative answer simply carries the coding forward from the previous month, thus improving accuracy.

The dependent interviewing facilitated by computer-assisted technology can also reduce respondent burden. For example, when a respondent is retired or on disability in one month, the interviewer knows of that status at later interviews and can check to see if the status is still the same. If it is, the person need not be put through all the labor force status questions.

Finally, for the important "looking for work" question, a much clearer distinction is made in the new questionnaire between "active" job search and "passive" job search. The response categories have been changed to provide a list of possible responses that distinguishes between active and passive activity. (See numbers 8 and 9 in the Revised Questionnaire col. and numbers 5 and 6 in the Unrevised CPS col. of table 1). This change will reduce errors generated when interviewers mistakenly recorded such passive job-search strategies as merely reading the want ads. In such cases, the respondents would have been categorized as looking for work and, hence, mistakenly considered unemployed. In addition, the change should also provide greater information on what unemployed workers actually do to try to find work.

Thus, the new CPS questionnaire should more accurately reflect respondent experience, provide more information for labor market analysis, and be more user-friendly and easier to process. In addition, the use of CAPI in the first and fifth visits and of CATI in other interviews should provide a data base that is amenable to obtaining faster answers to labor market analytical questions.

Effect on the Data

ESTIMATES FOR THE COUNTRY AS A WHOLE

In mid-November 1993, the Bureau of Labor Statistics announced the first results from the overlap sample with 12,000 households using the new questionnaire and computerized collection techniques. At that time, data from the overlap survey for the 12 months from September 1992 through August 1993 were analyzed in a series of conferences with users. Further, a series of articles explaining the need for, mechanics of, and effects of change was published (see Bowie, Calhoon, and Martin 1993; Bregger and Dippo 1993; Polivka and Rothgeb 1993a). The new survey's overlap sample showed a dramatic increase in labor force activity for adult women, a decline in adult men's labor force participation, and some increase in labor force activity for teenagers and for older workers. The changes resulted in a much higher overall unemployment rate than the old survey had shown. The test results show a jobless rate of 7.6 percent, a full five-tenths higher than in the current survey (table 2). The jobless rate difference between the overlap sample and the unrevised CPS was especially large for adult women—0.8 percentage point. The following discussion of differences is based on BLS conference materials.

These differences in labor force activity and in unemployment seem to be the result of the change in the very first set of labor force questions. The old question (which asked what the respondent was doing most of last week and was followed by several examples) seems to have produced a significant gender bias, partly because of the wording of the question and partly because of the manner in which interviewers tailored the question to individual respondents. The examples following the initial question were used to ask male respondents if they were "working or something else?" and to ask those who appeared to be homemakers if they were "keeping house or something else?" As a result, a number of women apparently were misclassified as not in the labor force. The wording of the initial question on the old questionnaire may also have led some respondents to indicate that they were employed, since work is socially desirable.

Several other changes in the questionnaire may also have affected the count of the unemployed. Respondent confusion in interpreting the old questions about layoff probably missed some persons who had been laid off and may have misclassified others, and the changes in the "looking for work" question could well have changed the number of people classified as unemployed in the overlap survey.

Although BLS and Census staffers had been concerned for some time about the possible gender bias in the first question, the fact that so much of the increased measured joblessness was among adult women was unexpected. The increase in joblessness among adult women also probably more accurately reflects the increase in the part-time and intermittent labor force.

Discouraged workers are not included in the unemployment count because they did not seek work during the 4 weeks preceding the survey, but an increase in the number of discouraged workers is often

Table 2. Rates and Ratios by Demographic Groups (Averages Based on September 1992–August 1993)

	Labor Force Participation Rate			Employment-to-Population Ratio			Unemployment Rate		
	Parallel Survey	CPS	Difference	Parallel Survey	CPS	Difference	Parallel Survey	CPS	Difference
All workers	66.7	66.2	.5*	61.7	61.5	.2	7.6	7.1	.5*
Adult men	76.5	77.0	−.5*	71.3	71.9	−.6*	6.9	6.7	.2
Adult women	59.6	58.3	1.3*	55.5	54.8	.7*	6.8	6.0	.8*
Teenagers	54.1	51.7	2.4*	43.1	41.6	1.6*	20.3	19.7	.6
White males	75.9	76.2	−.3	70.8	71.2	−.4	6.7	6.5	.2
White females	59.5	57.8	1.7*	55.6	54.5	1.1*	6.5	5.8	.7*
Black males	67.2	69.2	−1.9*	56.8	59.1	−2.2*	15.5	14.6	.9
Black females	57.8	57.5	.3	49.6	50.2	−.7	14.2	12.7	1.6

SOURCE.—U.S. Bureau of Labor Statistics 1993.

NOTE.—CPS denotes estimates from the Current Population Survey (source of official estimates). Differences may not be exact due to rounding.

* Statistically significant at a 10 percent level.

thought to be related to a rise in economic hardship and the number is, therefore, watched carefully. In the new questionnaire, a respondent is classified as a discouraged worker only if he or she has looked for work in the last year and is available to take a job if one were offered (table 3). This change in definition was recommended by the Levitan Commission to make the category more objective. The change had a substantial effect on the number of discouraged workers. The overlap sample data covering the September 1992–August 1993 period estimated only 424,000 discouraged workers, compared with 1,115,000 estimated by the current CPS.

The third group that is especially important in evaluating the extent of underemployment in the country is the number of persons who are employed part-time for economic reasons. This group, those who work part-time but who want full-time work, are sometimes called the partially unemployed. The new questionnaire has specific questions on whether the worker wants to and is available to work full-time rather than inferring the status from the reasons provided for part-time work. The number of people counted as working part-time for economic reasons was much smaller in the overlap survey than in the old CPS. The overlap results showed 4.9 million in this category, some 1.4 million fewer workers than in the unrevised CPS.

ESTIMATES AT THE STATE LEVEL

Because funds were limited, the overlap sample, although large by nongovernmental survey standards, is small compared with the 60,000 household monthly CPS surveys. That small size, and the fact that it was designed to represent the country as a whole rather than to provide data for individual states, means that changes at the state level have less reliability than those at the national level. Emphasizing the limited confidence in the statistical reliability of differences below the national level, the BLS nevertheless released estimates of the differences between the unrevised CPS and the overlap sample for large regions of the country and for several of the largest states.[3]

Regional unemployment rates, like the national rate, were higher in the new survey sample than in the old CPS. The unemployment rate for the 11 direct-use states as a whole was 0.5 percentage points higher in the new overlap survey than in the old. Unemployment rate differences between the overlap survey and the unrevised CPS ranged from no difference for California to a 1.2 percentage point difference for New York.

[3] It should be noted that in the past, monthly estimates for 11 large states were produced directly each month from the CPS because the states had large enough samples to permit compilation and publication of the data.

Table 3. Discouraged Workers Questions

Parallel Survey	Current Population Survey
1. Do you currently want a job, either full- or part-time? [If "yes" or "maybe," ask 2.]	1. Do you want a regular job now, either full- or part-time? [If "yes," ask 2.]
2. What is the main reason you were not looking for work during the LAST 4 WEEKS?	2. What are the reasons you are not looking for work? [Mark each reason mentioned.]
3. Did you look for work at any time during the last 12 months? [If "yes," ask 4.]	
4. Did you actually WORK at a job or business during the last 12 months? [If "yes," ask 5; otherwise, ask 7.]	
5. Did you do any of this work during the last 4 weeks? [If "no," ask 6.]	
6. And since you LEFT that job or business, have you looked for work? [If "yes," ask 7.]	
7. LAST WEEK, could you have started a job if one had been offered? [If "no," ask 8.]	
8. Why is that?	
To be classified as discouraged, an individual has to want a job, provide a discouraged reason for not looking, have looked in the last year (or since last working), and be available to work (where one's own temporary illness does not disqualify an individual from being available).	To be classified as discouraged, an individual has to want a job and provide only discouraged reasons for not looking.

SOURCE.—U.S. Bureau of Labor Statistics 1993.

HOW WILL THE DATA CHANGE?

These estimates of the differences between the old and the new questionnaires and collection procedures present the best general guide to the changes that will occur in the new Current Population Survey, but the exact month-to-month differences are difficult to predict. The half percentage point increase in the unemployment rate discussed above is, as we noted, based on the 12-month average from the overlap sample from September 1992 through August 1993. During that 12-month period, however, variability occurred from one month to the next—1-month estimates of the increase ranged from one-tenth of a percent to a full percentage point. We must also recognize that when all 1,500 regular CPS interviewers start using the new questionnaire and computer technology, some interviewer effects may occur, even though the Census Bureau has provided extensive training to these interviewers in preparation for the shift. In addition, seasonal patterns computed from the old questionnaire will be used in January 1994 to adjust the new survey to account for seasonal variation in the data. This procedure will continue for some time since 3–5 years of data are necessary to estimate new seasonal factors.

Another technical problem is that in January, the composite estimator designed to reduce month-to-month variations in the survey will not be used. The composite estimator in any month gives somewhat more weight to current-month rotation groups also in the survey from the previous month than to new rotation groups. But, in January, all households will respond to the new questionnaire for the first time, thus making it impossible to use compositing. The Bureau of Labor Statistics has announced that compositing will begin in April. The fact that the new questionnaire asks questions about discouraged workers each month rather than only in the month a household leaves the survey (months 4 and 8) could introduce further differences in the data from each of the rotation groups.

Introduction of Adjusted 1990-based Population Controls

After each decennial Census, new population controls are introduced into the CPS to weight the survey estimates to represent the population as a whole. This introduction will be done at the beginning of 1994 concurrent with the other improvements in the redesign. In this case, the Bureau of Labor Statistics, reflecting its goal of using all available research to improve the labor force estimates, decided to use the 1990 census counts adjusted for the estimated undercount of the population. The adjusted estimates should improve the estimates for some states and for minority groups.

The introduction of the 1990-based controls, adjusted for the undercount, will affect the levels of labor force and employment. The BLS has estimated that the labor force may increase by a little more than a million people as a result of the new controls, and total employment may rise by nearly 900,000. This is primarily because of the increased population of minorities.

The population control change, however, is also expected to increase the national unemployment rate by an additional tenth of a percentage point. When added to the 0.5 point increase anticipated from the questionnaire and technology changes in the redesign, therefore, the new CPS is expected to register an unemployment rate that is about six-tenths of a point higher than that measured by the unrevised survey.

Conclusion

The redesign of the CPS has been a difficult and expensive undertaking. It is a redesign that has incorporated state-of-the-art techniques in a survey of tremendous importance to the understanding of social and economic conditions. The two government agencies involved in the redesign have demonstrated not only their cooperation in technical competence but also their courage to move forward to bring about improvements. Changing a major economic indicator, especially one with the political sensitivity of the labor force survey, is always hard to accomplish. Improvement always brings risk, and assessment of that risk is not always easy. But retaining the status quo does not eliminate risk either. There is always the danger that the data produced from a static survey will increasingly lose relevance in a world where change is constant—change in labor force behavior, change in the meaning of words, and change in the way in which people think about the questions that are asked.

This tension between change and stability is especially relevant to surveys like the CPS, which produces time-series information. The extensive revamping of the survey questionnaire and the new collection techniques introduced in this redesign may produce abrupt changes, especially in the first few months, that make analysis of the data extraordinarily difficult. The differences between the old and the new surveys may cause problems for those who rely on the unemployment rate to allocate federal funds to states and local areas, causing them to ask whether the changes were necessary. How will benefits under the unemployment insurance laws be extended further? How will training funds under the Joint Training Partnership Act be allocated? How will decisions be made for the provision of emergency food and shelter?

Many Americans who do not understand survey design will criticize the government agencies for having caused confusion in understanding the labor market during the transition period. Others will try to use the increased unemployment resulting from the survey changes for their own political advantage. And still others may lose trust in the statistics produced by a government that admits that it has probably underestimated unemployment for many years.

Life will not be easy at the Bureau of Labor Statistics, which has the responsibility to explain and analyze the data from the new survey. A great deal of staff time will be required to answer inquiries and to brief the press, the professional community, and the public in general about the new survey. Statistical techniques will have to be developed to produce a bridge between the old and the new data series so that analysts can continue to use them. Changes may have to be made in the level of unemployment rates used to trigger on or off a number of individual government programs.

But so long as these efforts are conducted in an open and service-oriented manner, public understanding can be developed. The 1990 redesign of the CPS is an example of the kind of work that the two largest statistical agencies in the government, when working together for scientific excellence, can accomplish. A good statistical system must be alert to the need for continuous improvement of its data because statistical agencies that do not move forward will surely fall behind. The CPS redesign effort is one major effort to move forward.

References

Bowie, C. E., L. E. Calhoon, and E. Martin. 1993. "Overhauling the Current Population Survey: Evaluating Changes in the Estimates." *Monthly Labor Review* 116:29–33.

Bregger, J., and C. S. Dippo. 1993. "Overhauling the Current Population Survey: Why Is It Necessary to Change"? *Monthly Labor Review* 116:3–9.

Cannell, C., L. Oksenberg, F. J. Fowler, G. Kalton, and K. Bischoping. 1989. "New Techniques for Pretesting Survey Questions." Final Report for Grant Number HS 05616 from the National Center for Health Services Research and Health Care Technology Assessment. Ann Arbor: University of Michigan, Survey Research Center.

Dippo, Cathryn, and Janet L. Norwood. 1992. "A Review of Research at the Bureau of Labor Statistics." In *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*, ed. Judith M. Tanur, pp. 271–90. New York: Russell Sage.

Esposito, J. L., J. Rothgeb, A. E. Polivka, J. Hess, and P. C. Campanelli. 1992. "Methodologies for Evaluating Survey Questions: Some Lessons from the Redesign of the Current Population Survey." Paper presented at the International Conference on Social Science Methodology, Trento, Italy, June 22–26.

Fracasso, M. P. 1989. "Reliability and Validity of Response Categories for Open-ended Questions in the Current Population Survey." In *Proceedings of the Section on Survey Research Methods*, pp. 481–85. Alexandria, VA: American Statistical Association.

Grais, Bernard. 1979. "Classification of Persons Laid Off in Employment and Unemployment Statistics: Working Party on Employment and Unemployment Statistics." Document MAS/WP 7(79)5. OECD, Paris.

Jabine, Thomas B., Miron L. Straf, Judith M. Tanur, and Roger Tourangeau. 1984. *Cognitive Aspects of Surveys: Building a Bridge between Disciplines*. Washington, DC: National Academy Press.

National Commission on Employment and Unemployment Statistics. 1979. *Counting the Labor Force*. Washington, DC: Government Printing Office.

Palmisano, M. 1989. "Respondent Understanding of Key Labor Force Concepts Used in the CPS." Paper presented at the annual meeting of the American Statistical Association, Washington, DC, August.

Polivka, A. E., and J. M. Rothgeb. 1993a. "Overhauling the Current Population Survey: Redesigning the Questionnaire." *Monthly Labor Review* 116:10–28.

———. 1993b. "Redesigning the Questionnaire for the Current Population Survey." Paper presented at the annual meeting of the American Economic Association, Anaheim, CA, January 5–7.

President's Committee to Appraise Employment and Unemployment Statistics. 1962. *Measuring Employment and Unemployment*. Washington, DC: Government Printing Office.

U.S. Bureau of Labor Statistics. 1993. Conference Briefing Materials, Washington, DC.

U.S. Bureau of Labor Statistics and U.S. Census Bureau. 1986. Unpublished report, Bureau of Labor Statistics–Census Bureau Questionnaire Design Task Force, Washington, DC.

THE POLLS—POLL TRENDS
NUCLEAR POWER: THREE DECADES OF PUBLIC OPINION

EUGENE A. ROSA
RILEY E. DUNLAP

The last *Public Opinion Quarterly* report on nuclear power appeared in the aftermath of the April 1986 explosion of the nuclear power plant at Chernobyl in the Ukraine region of the former Soviet Union (de Boer and Catsburg 1988). Public support for nuclear power declined and concerns about nuclear safety increased in virtually all polls taken immediately after the accident. Surveys a year or so later, however, showed signs that nuclear power was regaining some of its lost ground, leading de Boer and Catsburg to hypothesize that "these large changes [toward increased opposition] in public opinion are likely to be temporary" (1988, p. 255). This "rebound" hypothesis echoed the earlier observations of other analysts who noted an upturn in support for nuclear power in the months following the 1979 Three Mile Island (TMI) accident in Pennsylvania (Freudenburg and Rosa 1984). Long-term survey results, while somewhat mixed, seriously challenge the rebound hypothesis, however, as trends in the longest running time-series items have become increasingly unfavorable to nuclear power.

Long-Term General Trends

The longest running time series concentrate on attitudes toward the continued growth of nuclear power. Question 1 presents the responses

---

EUGENE A. ROSA is professor of sociology and RILEY E. DUNLAP is professor of sociology and rural sociology at Washington State University. For help with data compilation the authors thank Barbara C. Farhar of the National Renewable Energy Laboratory of the U.S. Department of Energy, Sue Dodd of the Louis Harris Data Center at the University of North Carolina, Robert Persons of the Roper Center for Public Opinion Research at the University of Connecticut, Diane Chess of Media General, Marilyn Reed and Keating Holland of Yankelovich Partners, Inc., and Ed Keller and John Cianelli at the Roper Organization, Inc. Special thanks are due Ann Stouffer Bisconti, vice president for research and program evaluation of the U.S. Council for Energy Awareness, who kindly shared all the nuclear opinion data sponsored by her organization, and to Donna Poire who provided able typing.

Public Opinion Quarterly Volume 58:295–325 © 1994 by The American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5802-0001$02.50.

to a Harris item from 1975 to 1990 asking about building more nuclear plants. Until the late 1970s a majority of Americans supported, often by wide margins, growth in nuclear power (an exception is the April 1976 result showing a plurality of support). Immediately following the accident at Unit 2 at Three Mile Island in Pennsylvania in March 1979, this support eroded considerably but still consistently exceeded opposition. With the May 1979 survey, support had regained its majority position but appeared to be standing on much shakier ground than before the accident: over the course of a dozen surveys conducted over the next 22 months support for nuclear power, though relatively consistent, vacillated between bare majority and plurality. In one case—May 1980—opposition actually exceeded support, by 10 percentage points, but this was an apparent anomaly.

In contrast, the October 1981 result showing majority opposition signaled a break with the past, indeed, a reversal in public preference for nuclear growth. Since then, opposition has consistently exceeded support. In fact, over the last four surveys spanning 7 years, opposition has exceeded support by a margin of about 2:1, a complete flipflop from the earliest Harris survey.

In the fourth quarter of 1974, Cambridge began asking Americans their views about building more nuclear power plants as a way of dealing with the energy crisis. Cambridge contextualized their question within a concern for the energy crisis because they first polled the public on the subject in the months following the Arab oil embargo and the emerging "energy crisis" (Rosa 1978).

Question 2 presents the Cambridge results, which, similar to Harris, show that nuclear power enjoyed consistently strong support throughout the 1970s, at least up until the TMI accident in March 1979. The effect of the accident on public attitudes is apparent, as opposition exceeded support for the first time in the second quarter 1979 survey. By the third quarter of that year, however, support had rebounded sufficiently to outpace opposition but not by as large a margin as before the accident. Further, while support remained consistent over the next 2.5 years, it was only a plurality of support, and sometimes barely that. Then, similar to the Harris trend, a sharp reversal in opinion appeared (showing up for Cambridge in the first quarter of 1982). What is most noteworthy about this shift from majority support to majority opposition to further nuclear plant construction is that it endured across subsequent surveys during the following 8 years. For most surveys the ratio of opposition to support approaches or exceeds 2:1—again, a near-perfect reversal of public preference in the early 1970s.[1]

---

[1] Using similar but somewhat different wording from Cambridge, NBC News posed the following to respondents in August 1990: "Let me read you a number of proposals

A second Cambridge item asked about the construction of more nuclear power plants, but without any mention of an energy crisis. Question 3 reveals an overall pattern similar to that of the Harris and the first Cambridge item: consistent support for nuclear power eventually gave way to strong opposition. Despite this overall similarity among the three times series, there are a few anomalies. The spring 1975 results for the second Cambridge item, showing a plurality of 45 percent support for building more nuclear power plants, is substantially lower than the Harris (63 percent) and the first Cambridge item (54 percent). Similar discrepancies showing noticeably lower support with the second Cambridge item are found in July 1976 and January 1977. While the reason for these anomalies is not entirely clear, part of the answer may be traced to the fact that until 1981, when levels of support converge for all three series, the second Cambridge item elicited a higher proportion (between 18 percent and 32 percent) of "don't know" responses than the other two items.

During the early 1980s, as the American public was apparently reevaluating nuclear technology, NBC News asked about the building of more nuclear power plants (question 4). All four results show majority opposition, the last two by the same wide margins evident in the Harris and Cambridge results. Later Yankelovich polls[2] in 1991 and 1992 (question 5) show majority opposition (combining the "somewhat" and the "strongly" categories), with the more recent one showing greater than a 2:1 ratio of opposition to support for more nuclear power plant construction.

A consistent pattern of growing opposition to the building of more nuclear power plants emerges from these results. The pattern comprises three distinct stages: an early stage in the 1970s when Americans were enthusiastic about the growth of nuclear power; a second stage of ambivalence following TMI when a less enthusiastic plurality of citizens consistently supported nuclear growth; and a third stage—emerging in the early 1980s—when a decisive majority of Americans opposed building more nuclear power plants. Apace with the decline in support for nuclear power plants was the decline in the proportion

---

some people have suggested that America should pass in order to avoid an energy shortage. For each, please tell me if you favor or oppose that suggestion. . . . Encourage construction of more nuclear power plants." A solid 57 percent majority opposed the proposal, while 37 percent favored it. While the ratio of opposition to support was not as pronounced as the 2:1 Cambridge results, the solid majority opposition is consistent with the general picture produced by both the Cambridge and Harris series.

[2] Also in March 1992 Frederick/Schneiders polled 750 Americans with the question: "Do you favor or oppose the building of more nuclear power plants in this country?" The 65 percent who said they were opposed was virtually identical to the 64 percent opposed (44 percent strongly opposed and 22 percent somewhat opposed) in the March 1992 Yankelovich poll.

of "don't know's," reflecting, perhaps, a crystallization of nuclear attitudes.

The evolution in opinion about nuclear plant construction, including opinion reversal, can be best seen in the trend lines in figures 1 and 2 for the longest running—Harris (question 1) and Cambridge (question 2)—data series. Most striking in both figures are the impact of the TMI accident on public attitudes, the abrupt reversal in attitudes in the early 1980s, and the steady growth in opposition following the reversal.

The data also show that the reversal in attitudes, with opposition overtaking support, occurred over 4 years before the Chernobyl accident. This may be a consequence of growing concern at the time about nuclear weapons and nuclear war. The NATO resolution in late 1979 to deploy medium-range Cruise and Pershing II missiles signaled a demise in detente and fueled renewed worries about nuclear war. As a result, the peace movement attracted a large enough following to become a mass movement, leading to large-scale peace and antinuclear demonstrations in many NATO countries (de Boer 1985).

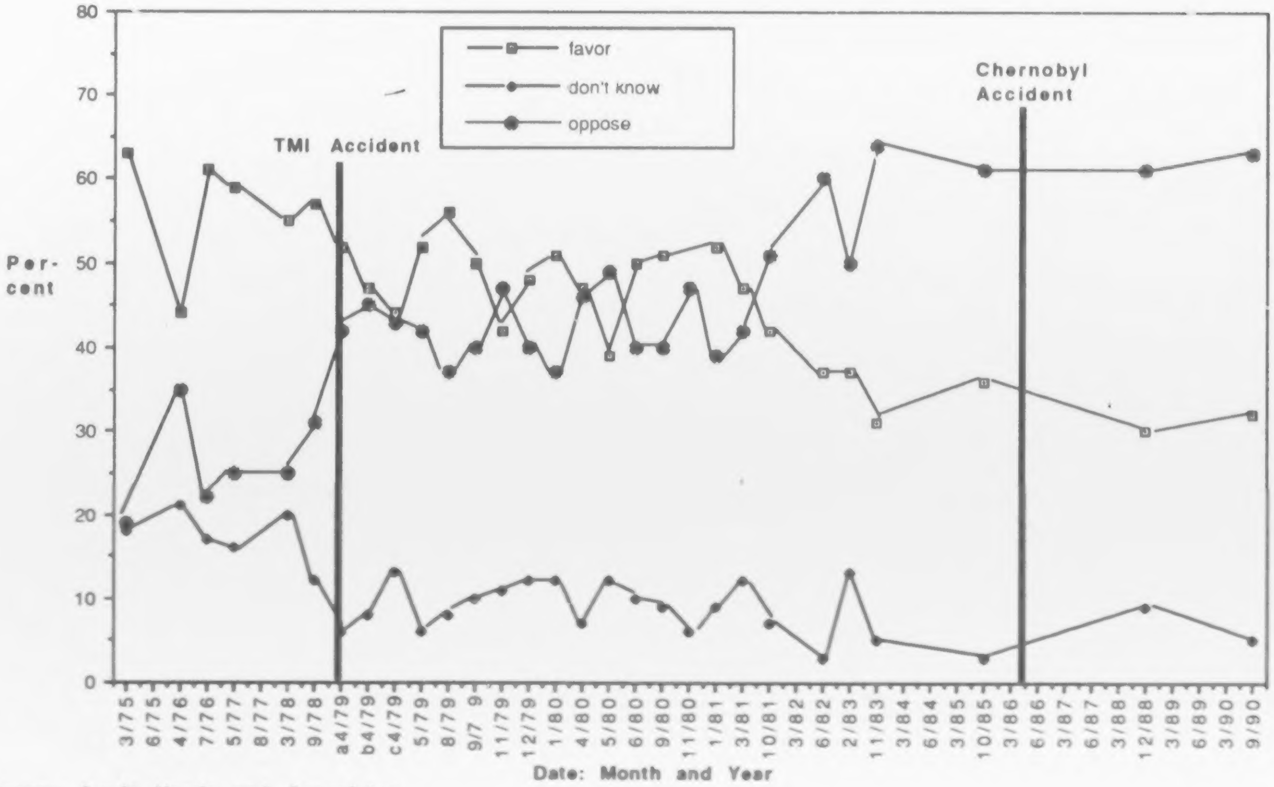
In 1981 the Reagan administration ordered the production of the neutron bomb, previously abandoned by President Carter, and continued to deploy missiles in Western Europe. As a result, American nuclear activists shifted their efforts from nuclear power plants to nuclear weapons (Nelkin 1981). During 1981 and 1982, as the nuclear freeze gained momentum, there was an upsurge in fear of nuclear war among Americans—whether measured by public opinion polls (Smith 1988), by the volume of survey questions on nuclear weapons/war items (Kramer, Kalick, and Milburn 1983), or by the volume of media coverage (Mazur 1990).

This anxiety about nuclear weapons peaked in early 1982, at precisely the time of the reversal in attitudes toward constructing more nuclear power plants, suggesting that weapon worries had spilled over to the domain of commercial nuclear power. The spillover appears to have initiated a steady growth in opposition to nuclear plant construction that continued to the very end of our data series. Interestingly, contrary to a widely held view, the 1986 Chernobyl accident did not have a marked impact on public attitudes: at the time of the accident opposition already exceeded support by nearly 2:1 (see fig. 2).

Clearly the American public no longer supports the construction of additional nuclear power plants. Despite the fact that items asking about nuclear plant construction are widely used, questions have been raised about their validity as a global measure of nuclear attitudes (Bisconti 1988). Responses to them should be interpreted in the context of other items.[3]

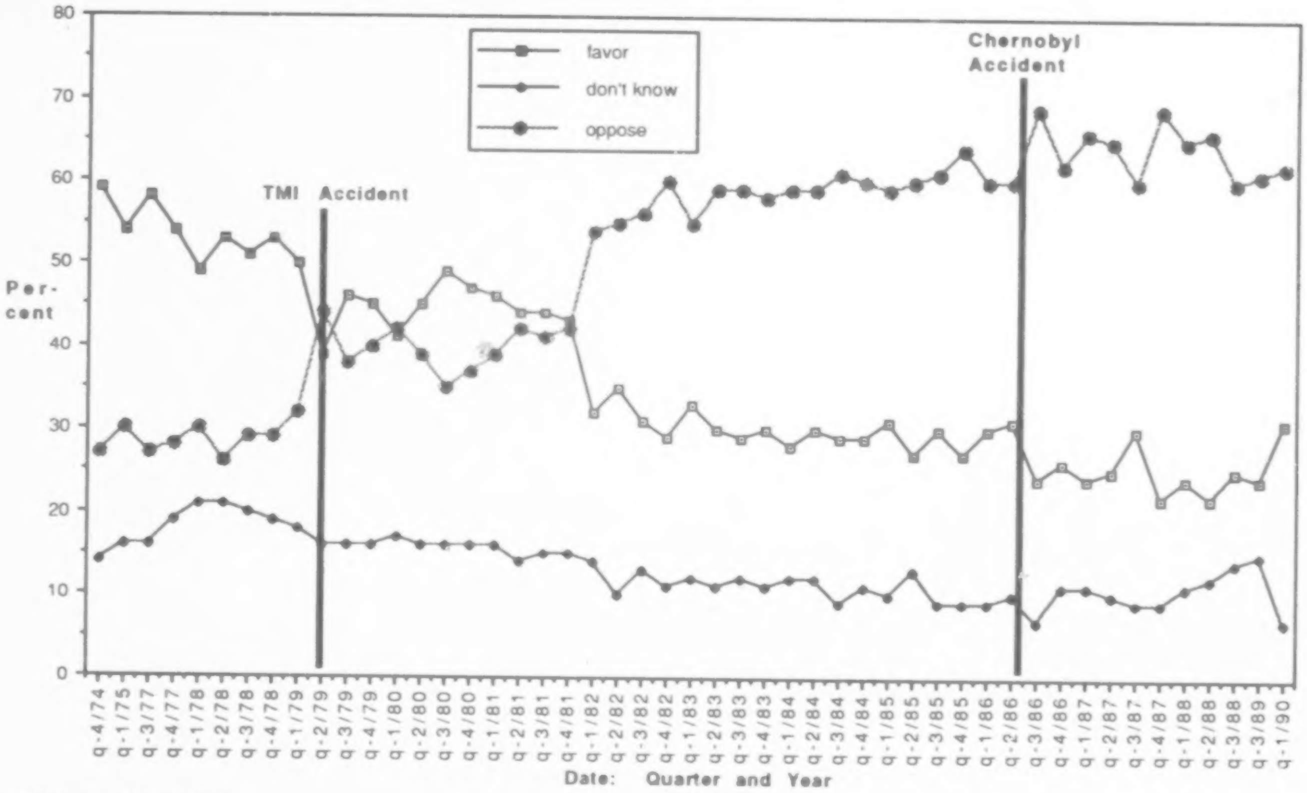
---

[3] Bisconti (1991) has argued that the item is a useful gauge of public attitude only when considered in proper context. Responses to the item are presumed to reflect public belief



Source: Louis Harris and Associates
National Samples

Figure 1. Public attitudes toward building nuclear power plants in the United States



Source: **Cambridge**
National Samples

Figure 2. Public attitudes toward building nuclear power plants in the United States as a way of dealing with the energy crisis.

Trends by Topic Area

NUCLEAR POWER AS A SOURCE OF ELECTRICITY

Despite the long-term trends showing a growing opposition to constructing more nuclear power plants, perhaps the public is more supportive of the technology when it is connected directly to its instrumental purpose: the generation of electricity, which is the sole commercial use of nuclear power. In an Associated Press/Media General poll shortly after the Chernobyl accident the public was evenly split on the use of nuclear power to generate electricity, but in 1989 a strong 55 percent majority favored it (question 6). For all three Bruskin and Goldring polls conducted between 1992 and 1993 a majority of respondents favored (strongly or somewhat) the use of nuclear power to generate electricity (question 7). It should be emphasized, however, that these polls did not ask for unqualified support for nuclear power, but only as one way of providing electricity.[4] As for the present, Americans are apparently aware that nuclear is already an important energy source for the generation of electricity. In Cambridge surveys in 1986 and 1989, a sizable majority, 70 percent, agreed that nuclear power was important in meeting the nation's electricity needs today (question 8).

LOCAL SITING OF POWER PLANTS

It is one thing to have an opinion about the construction of nuclear power plants in the abstract; it is another to be confronted with the prospect of having a plant built nearby. As early as 1976 Gallup found that a slight plurality of respondents opposed the construction of a nuclear power plant within 5 miles of where they lived; 10 years later opposition reached a whopping 73 percent (question 9). It might be expected that Americans would be more tolerant of local nuclear plants if they were reminded that the purpose of the technology was to pro-

---

in the need for additional electricity-generating plants. Recent polls, such as Frederick/Schneiders, Inc., in March 1992 where only 28 percent thought the United States needed more power plants, indicate the public remains unconvinced of the need for new electrical generating capacity. Once Americans recognize the need for increased electricity supply, the argument goes, there will be a revival in support for nuclear power.

[4] Support for nuclear power as an electricity source may be due to a lingering belief in the possibility of an electricity shortage. In 1990 Cambridge asked: "How likely do you think it is that we could have serious electricity shortages in the United States in the next few years?" A majority indicated that shortages were likely, 22 percent saying "very likely" and 31 percent saying "likely" (30 percent "not too likely," 11 percent "not at all likely," and 6 percent didn't know).

vide electricity. Cambridge and Gallup, using an identically worded item, assessed this possibility in a series of polls from 1986 to 1990 (question 10). The most popular of three possible responses, attracting near-majority to clear-majority support, was to "reserve judgment"—perhaps reflecting, on the one hand, a reluctance to reach a firm position until apprised of the specific circumstances surrounding such a decision and, on the other hand, a willingness to leave open the nuclear option as a hedge against future electricity shortages. Of those who had evidently formed an opinion about a local plant, opposition consistently outweighed support, but by a declining margin over time.

PRAGMATISM

Does the sizable opposition to the construction of more nuclear power plants translate into a deep-seated aversion to the technology, into an unreflective rejection of all matters nuclear, or even a willingness to get rid of the technology entirely? Or is the opposition the result of sensible caution about a large-scale technology that is neither simple nor risk free? Cambridge and Gallup asked respondents (question 11) whether nuclear power was a good choice, a realistic choice, or a bad choice. Strong pluralities to majorities consistently said it was a realistic choice. The remaining 40–50 percent of the opinionated public were more likely to say it was a bad rather than good choice, but sometimes by only narrow margins.

As for eliminating nuclear power, only a minority of Americans in a Cambridge item (question 12) asked between 1983 and 1991 have supported this option. The most frequently chosen option, attracting sizable pluralities to majorities, favored the status quo: let the existing nuclear plants operate but do not build any more. A similar finding emerges from a Bruskin and Goldring poll (question 13): again, the middle category favoring the status quo attracted near-majorities to majorities of respondents. Taken together, the picture that emerges from these items is not one of a knee-jerk public passionately committed to the riddance of nuclear energy but of a pragmatic public weighing technological risks against unforeseeable future energy needs. As a hedge against a shortfall in future energy supply, Americans appear willing to leave the nuclear option open.

NUCLEAR FUTURES

What does the public believe the future holds for nuclear power? In 1989 Cambridge began to ask Americans (followed up by Gallup and Bruskin and Goldring with the identical item) how important a role, practically speaking, nuclear power should play in America's energy future (question 14). Between two-thirds and three-quarters of Americans thought nuclear power should play either a very important or somewhat important future role.

When it comes to the public's expectations about the future of nuclear power, instead of its preference for growth, a clear and consistent picture emerges. Four separate polls have asked Americans to assess the importance of nuclear power in meeting future energy needs. Despite considerable differences in question wording among the polls (questions 15, 16, 17, and 18), a steady three-quarters of Americans believe that nuclear power plants will be built in the future and that the importance of nuclear as an energy source will increase.

NUCLEAR COMPARED TO OTHER ENERGY SOURCES

The results from the two previous sections present a paradox. On the one hand, Americans are solidly opposed to the construction of more nuclear power plants, especially opposed to their local siting; on the other, they solidly favor leaving the nuclear option open and firmly believe nuclear power will play an important role in the nation's electricity future. The paradox may reside in the fact that two separate cognitive domains may have been tapped by the different types of questions on the topic. Thus, the time-series questions asking about nuclear construction and siting may be tapping attitudes toward the technology—the public's feelings and preferences about nuclear power growth. The questions with a future orientation may be tapping beliefs about this source of energy—what the public believes the actual nuclear future is likely to look like.

In two recent polls Yankelovich (question 19) has asked respondents to consider nuclear power in comparison with coal and oil. Near majorities have said that we should rely less on nuclear power in the future, but they have said the same—even more emphatically—about the other two choices. Another comparative perspective was provided by Roper (question 20) when it asked respondents to compare nuclear power, natural gas, coal, and oil on nine attributes. Nuclear power was generally seen to be better for the environment than coal or oil, more economical than oil, and the energy-to-electricity source the nation is least likely to run out of. These positive perceptions of nuclear power, however, were more than counterbalanced by the vast majorities who saw it as the least safe electricity source, least acceptable for widespread use, and potentially most dangerous to human health.

Another question comparing nuclear to other electricity producing energy sources was asked by Opinion Research Corporation (question

21), which focused on the energy source most responsible for pollution. While nuclear power attracted majority but declining beliefs that it was the least polluting source for generating electricity in the first two polls in 1976 and 1980, its rating dropped farther from 1983 to 1989 (based on the same question but with natural gas included as a response option). Nuclear power's image as a fairly clean energy source has been on the decline.

RISK AND DANGER

Lying behind people's perceptions of energy alternatives are issues concerning the levels of risk and danger associated with them (Farhar 1993; Gould et al. 1988), and these concerns can affect the responses to nuclear opinion items. For example, Mitchell (1980a, p. 7) concluded from earlier polls that opposition or support for nuclear power "can be made to shift 40 percentage points, just by varying the assurance of safety" within the question wording. We have already seen in Roper's comparative question (question 20) that the public views nuclear power as unsafe and potentially dangerous to human life. Further, Americans have increasingly said that a local nuclear power plant would present dangers (question 22), with the most recent polls finding two-thirds or more endorsing this position. As for the issue of risk (question 23), three-fourths have said that living near a nuclear plant involves moderate to high risk. Clearly, Americans are worried about the uncertainties associated with nuclear power.

NUCLEAR AS A TRADE-OFF CHOICE

Nearly all of imported oil is used for petroleum products, not for the generation of electricity (Energy Information Administration 1993). Nevertheless, Americans have been asked whether they would favor nuclear power if it would reduce the nation's dependence on foreign oil. Presented with such a trade-off, majorities of the public favor nuclear power—often by sizable margins (questions 24 and 25).

Recently scientists have pointed to a possible warming trend in the mean global temperature due, principally, to the burning of fossil fuels. Nuclear energy could substitute for the fossil fuels used to generate electricity and, accordingly, the public has responded to this trade-off: a majority of the public believes that using nuclear power will cut greenhouse gas emissions, and a decisive majority favors its use if it contributes to a reduction in greenhouse gas emissions and air pollution (questions 26a and 26b).

Conclusion

Taken together, the data show that, on the one hand, solid majorities of the public oppose the construction of more nuclear plants and are likewise opposed to their local siting; on the other, equally solid majorities believe that nuclear power should be and will be an important energy source in the nation's future. What accounts for the apparent contradiction? In our view the results do not reveal a contradiction but, rather, a pragmatic logic by which the public distinguishes between nuclear power in principle and in practice.

It appears that for nuclear power, as has been found for a broad range of issues (e.g., Page and Shapiro 1992), the public's collective preference for specific policy options can differ considerably from its general attitudes. Thus, Americans support the idea of leaving the nuclear option open, perhaps as a trump card against possible future energy shortages; but when it comes to the specific means for achieving that opinion—the siting and construction of nuclear power plants—they are solidly opposed.

What might this equivocation mean for the future of nuclear power? The answer will be shaped by the larger, decidedly unfriendly social climate regarding technological and environmental hazards. There is currently substantial concern about environmental quality (Dunlap and Scarce 1991), accompanied by growing media coverage and public awareness of a wide variety of societal risks (Russell 1988). Opposition to the local siting of nearly all environmental and technological hazards, evidenced by the well-known "not in my backyard"/NIMBY syndrome, is growing.[5] Because nuclear power is among the most dreaded technological risks (Slovic 1987) and because nuclear power plants are undesirable local facilities,[6] it is difficult to imagine a community accepting one (Rosa and Freudenburg 1993). The situation is exacerbated by widespread opposition to the siting of a permanent repository to store large and growing volumes of nuclear wastes (Dunlap, Kraft, and Rosa 1993). On balance, then, the public is apparently unwilling to rule out the nuclear option, but it is opposed to the implementation of this option in the near future.

---

[5] In the past decade, approximately 50 percent of attempts to site an oil refinery have failed due to public opposition and, with rare exception (e.g., the Pittsburgh and Denver airports), "no large metropolitan airport has been sited in the United States since the Dallas–Fort Worth facility in the early 1960s" (Shrader-Frechette 1991, p. 4). Similarly, few hazardous waste disposal facilities have been sited in the nation since 1980 because of public opposition (Kunreuther 1992).

[6] National surveys asking "tolerance" distances for various facilities find that nuclear power plants are the least or one of the least tolerable facilities (Mitchell 1980b), trailing only nuclear waste repositories and chemical-waste landfills (Slovic, Layman, and Flynn 1993).

All data reported in the tables that follow are based on national probability samples; sample sizes are provided whenever available.

Appendix

Sources and Abbreviations

B/G: Bruskin and Goldring
Cambridge: Cambridge Reports, Inc., and Cambridge Reports/Research International
F/S: Frederick/Schneiders, Inc.
Gallup: Gallup Organization
Harris: Louis Harris and Associates
AP/MG: Associated Press/Media General
NBC News: National Broadcasting Company
ORC: Opinion Research Corporation
Roper: The Roper Organization, Inc.
Yankelovich: Yankelovich, Clancy, and Shulman, and Yankelovich Partners, Inc.

References

Bisconti, Ann S. 1988. "Public Opinion." *Nuclear Industry* (July/August), pp. 7–9.

———. 1991. "Public Opinion on Nuclear Energy in the U.S.A. and Worldwide." Paper presented at the Sixth Annual Conference of the Korea Atomic Industrial Forum/Korean Nuclear Society, April 15–17.

de Boer, Connie. 1985. "The Polls: The European Peace Movement and Deployment of Nuclear Missiles." *Public Opinion Quarterly* 49:119–32.

de Boer, Connie, and Ineke Catsburg. 1988. "The Polls: The Impact of Nuclear Accidents on Attitudes toward Nuclear Energy." *Public Opinion Quarterly* 52:254–61.

Dunlap, Riley E., Michael E. Kraft, and Eugene A. Rosa, eds. 1993. *Public Reactions to Nuclear Waste: Citizens' Views of Repository Siting*. Durham, NC: Duke University Press.

Dunlap, Riley E., and Rik Scarce. 1991. "The Polls—Poll Trends: Environmental Problems and Protection." *Public Opinion Quarterly* 55:713–34.

Energy Information Administration. 1993. *Energy Facts 1991*. Washington, DC: Energy Information Administration/U.S. Department of Energy.

Farhar, Barbara C. 1993. *Trends in Public Perceptions and Preferences on Energy and Environmental Policy*. Golden, CO: National Renewable Energy Laboratory.

Freudenburg, William R., and Eugene A. Rosa. 1984. "Are the Masses Critical?" In *Public Reactions to Nuclear Power: Are There Critical Masses?* ed. William R. Freudenburg and Eugene A. Rosa, pp. 331–48. Boulder, CO: Westview/American Association for the Advancement of Science.

Gould, Leroy C., Gerald T. Gardner, Donald R. DeLucca, Adrian R. Tiemann, Leonard W. Doob, and Jan A. J. Stolwijk. 1988. *Perception of Technological Risks and Benefits*. New York: Russell Sage.

Kramer, Bernard M., S. Michael Kalick, and Michael A. Milburn. 1983. "Attitudes toward Nuclear Weapons and Nuclear War: 1945–1982." *Journal of Social Issues* 39:7–24.

Kunreuther, Howard. 1992. "A Conceptual Framework for Managing Low-Probability Events." In *Social Theories of Risk*, ed. Sheldon Krimsky and Dominic Golding, pp. 301–26. Westport, CT: Praeger.

Mazur, Allan. 1990. "Nuclear Power, Chemical Hazards, and the Quantity-of-Reporting Theory of Media Effects." *Minerva* 28:294–323.

Mitchell, Robert C. 1980a. *Final Results of the Resources for the Future National Environmental Survey for the President's Council on Environmental Quality*. Washington, DC: Resources for the Future.

———. 1980b. *Public Opinion on Environmental Issues: Results of a National Public Opinion Survey*. Washington, DC: Council on Environmental Quality.

Nelkin, Dorothy. 1981. "Anti-nuclear Connections." *Bulletin of the Atomic Scientists* 37:36–40.

Page, Benjamin I., and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.

Rosa, Eugene. 1978. "The Public and the Energy Problem." *Bulletin of the Atomic Scientists* 34:5–7.

Rosa, Eugene A., and William R. Freudenburg. 1993. "The Historical Development of Public Reactions to Nuclear Power: Implications for Nuclear Waste Policy." In Dunlap, Kraft, and Rosa 1993, pp. 32–63.

Russell, Christine. 1988. "Risk vs. Reality: How the Public Perceives Health Hazards." *Washington Post*, July 14, pp. 14–18.

Shrader-Frechette, Kristin S. 1991. *Risk and Rationality*. Berkeley and Los Angeles: University of California Press.

Slovic, Paul. 1987. "Perception of Risk." *Science* 236:280–85.

Slovic, Paul, Mark Layman, and James H. Flynn. 1993. "Perceived Risk, Trust, and Nuclear Waste: Lessons from Yucca Mountain." In Dunlap, Kraft, and Rosa 1993, pp. 64–86.

Smith, Tom W. 1988. "The Polls—a Report: Nuclear Anxiety." *Public Opinion Quarterly* 52:557–75.

Long-Term Trends in Attitudes toward Nuclear Plant Construction

1. In general, do you favor or oppose the building of more nuclear power plants in the United States?

			Harris			
	3/75 (%)	*4/76* (%)	*7/76* (%)	*5/77* (%)	*3/78* (%)	*9/78* (%)
Favor	63	44	61	59	55	57
Oppose	19	35	22	25	25	31
Don't know	18	21	17	16	20	12
	4/79[a] (%)	*4/79*[a] (%)	*4/79*[a] (%)	*5/79* (%)	*8/79* (%)	*9/79* (%)
Favor	52	47	44	52	56	50
Oppose	42	45	43	42	37	40
Don't know	6	8	13	6	8	10
	11/79 (%)	*12/79* (%)	*1/80* (%)	*4/80* (%)	*5/80* (%)	*6/80* (%)
Favor	42	48	51	47	39	50
Oppose	47	40	37	46	49	40
Don't know	11	12	12	7	12	10
	9/80 (%)	*11/80* (%)	*1/81* (%)	*3/81* (%)	*10/81* (%)	*6/82* (%)
Favor	51	47	52	47	42	37
Oppose	40	47	39	42	51	60
Don't know	9	6	9	12	7	3
	2/83 (%)	*11/83* (%)	*10/85* (%)	*12/88* (%)	*9/90* (%)	
Favor	37	31	36	30	32	
Oppose	50	64	61	61	63	
Don't know	13	5	3	9	5	

NOTE.—Sample sizes range from 1,200 to 1,600.

[a] Each is a separate survey in the same month.

2. I'm going to read you a list of proposals for dealing with the energy crisis, and I'd like you to tell me whether you generally favor or oppose each one . . . building more nuclear power plants. (Q = quarter.)

			Cambridge			
	Q4/74 (%)	*Q1/75* (%)	*Q3/77* (%)	*Q4/77* (%)	*Q1/78* (%)	*Q2/78* (%)
Favor	59	54	58	54	49	53
Oppose	27	30	27	28	30	26
Don't know	14	16	16	19	21	21
	Q3/78 (%)	*Q4/78* (%)	*Q1/79* (%)	*Q2/79* (%)	*Q3/79* (%)	*Q4/79* (%)
Favor	51	53	50	39	46	45
Oppose	29	29	32	44	38	40
Don't know	20	19	18	16	16	16

	Q1/80 (%)	*Q2/80* (%)	*Q3/80* (%)	*Q4/80* (%)	*Q1/81* (%)	*Q2/81* (%)
Favor	41	45	49	47	46	44
Oppose	42	39	35	37	39	42
Don't know	17	16	16	16	16	14

	Q3/81 (%)	*Q4/81* (%)	*Q1/82* (%)	*Q2/82* (%)	*Q3/82* (%)	*Q4/82* (%)
Favor	44	43	32	35	31	29
Oppose	41	42	54	55	56	60
Don't know	15	15	14	10	13	11

	Q1/83 (%)	*Q2/83* (%)	*Q3/83* (%)	*Q4/83* (%)	*Q1/84* (%)	*Q2/84* (%)
Favor	33	30	29	30	28	30
Oppose	55	59	59	58	59	59
Don't know	12	11	12	11	12	12

	Q3/84 (%)	*Q4/84* (%)	*Q1/85*[a] (%)	*Q2/85*[b] (%)	*Q3/85*[a] (%)	*Q4/85*[a] (%)
Favor	29	29	31	27	30	27
Oppose	61	60	59	60	61	64
Don't know	9	11	10	13	9	9

	Q1/86[a] (%)	*Q2/86*[b] (%)	*Q3/86*[a] (%)	*Q4/86*[c] (%)	*Q1/87*[c] (%)	*Q2/87*[d] (%)
Favor	30	31	24	26	24	25
Oppose	60	60	69	62	66	65
Don't know	9	10	7	11	11	10

	Q3/87[c] (%)	*Q4/87*[a] (%)	*Q1/88*[d] (%)	*Q2/88*[d] (%)	*Q3/88*[d] (%)	*Q3/89*[d] (%)	*Q1/90*[d] (%)
Favor	30	22	24	22	25	24	31
Oppose	60	69	65	66	60	61	62
Don't know	9	9	11	12	14	15	7

NOTE.—Wording about proposals was modified to suit the number of alternatives paired with nuclear power. Four such modifications occurred and are flagged to the quarters in which they were used. Sample sizes of approximately 1,500 until fourth quarter 1989, and 1,250 thereafter.

[a] "I'm going to read you two proposals for dealing with the energy crisis, and I'd like you to tell me whether you generally favor or oppose each one. . . . Building nuclear power plants."

[b] "I'm going to read you a proposal for dealing with the energy crisis. Please tell me whether you generally favor or oppose building more nuclear power plants."

[c] "Please tell me whether you generally favor or oppose building more nuclear power plants as a way of dealing with the energy crisis."

[d] "I'm going to read you several proposals for dealing with the energy crisis, and I'd like you to tell me whether you generally favor or oppose each one. . . . Building more nuclear power plants."

3. Do you generally favor or oppose the construction of more nuclear power plants?

			Cambridge		
	4/75[a]	*10/75*	*1/76*	*4/76*	*7/76*
	(%)	(%)	(%)	(%)	(%)
Favor	45	46	47	49	48
Oppose	35	23	34	33	31
Don't know	21	32	19	18	21
	10/76	*1/77*	*1/81*[b]	*10/81*[b]	*7/87*[c]
	(%)	(%)	(%)	(%)	(%)
Favor	52	51	51	48	30
Oppose	28	30	41	43	60
Don't know	20	19	8	9	9

NOTE.—Sample sizes of approximately 1,500.

[a] The original wording of the question (asked in 4/75 only) included the preliminary statement: "Another possible solution to the energy crisis is nuclear power. Do you generally favor or oppose nuclear power plants?"

[b] For 1981, the question and response categories were expanded to: "How do you feel about building more nuclear power plants? In general, do you strongly favor, somewhat favor, somewhat oppose, or strongly oppose building more nuclear power plants?" In reporting the results, we combine the two categories of favorability and the two categories of opposition.

[c] In 1987, the question was expanded to: "Please tell me whether you favor or oppose building nuclear power plants as a way of dealing with the energy crisis."

4. Do you think that more nuclear power plants should be built in the United States, or do you think they should not be built?

		NBC News		
	11/81	*1/82*	*11/82*	*4/86*
	(%)	(%)	(%)	(%)
Should be built	32	37	29	27
Should not be built	56	53	63	65
Not sure	12	10	8	8
N	1,602	1,599	1,583	1,599

5. Do you strongly favor, favor somewhat, oppose somewhat, or strongly oppose building more nuclear power plants in this country?

	Yankelovich	
	4/91	*3/92*
	(%)	(%)
Favor strongly	18	12
Favor somewhat	22	16
Oppose somewhat	20	20
Oppose strongly	32	44
Not sure	8	8
N	1,000	1,400

Nuclear Power as a Source of Electricity

6. Do you support or oppose using nuclear power to generate electricity?

	AP/MG	
	6/86[a]	*1/89*
	(%)	(%)
Support	36	55
Do not support	35	34
No opinion	26	11
N	1,365	1,162

[a] Wording was: "Do you support the use of nuclear energy to generate electricity or not, or don't you have an opinion either way?"

7. Overall, do you strongly favor, somewhat favor, somewhat oppose, or strongly oppose the use of nuclear energy as one of the ways to provide electricity for the United States?

		B/G	
	2/92	*12/92*	*2/93*
	(%)	(%)	(%)
Strongly favor	25	20	22
Somewhat favor	35	30	31
Somewhat oppose	18	21	19
Strongly oppose	15	18	18
Don't know	7	11	11

NOTE.—Sample sizes of approximately 1,000.

8. How important do you think nuclear energy plants are in meeting this nation's electricity needs today? Do you think nuclear energy plants are very important, somewhat important, not too important, or not important at all?

	Cambridge	
	8/86	*8/90*
	(%)	(%)
Very important	32	34
Somewhat important	38	36
Not too important	16	13
Not important at all	12	12
Don't know	2	5

NOTE.—Sample size of approximately 1,500 for 1986, and 1,250 for 1990.

Local Siting of Power Plants

9. As of today, how do you feel about the construction of a nuclear power plant in this area—that is, within five miles from here? Would you be against the construction of such a plant in the area, or not?

		Gallup	
	1976	*1979*	*6/86*
	(%)	(%)	(%)
Against	45	60	73
Not against	42	33	23
No opinion	13	7	4

NOTE.—Sample sizes of approximately 1,000.

10. If a new power plant were needed to supply electricity *in your area*, would you favor a nuclear power plant, oppose a nuclear energy/power plant, or reserve judgment until you had more information? (Emphasis added.)

	Cambridge 11/86 (%)	*Cambridge 2/87* (%)	*Cambridge 5/87* (%)	*Cambridge 8/87* (%)	*Gallup 9–10/87* (%)
Favor	18	15	14	17	19
Oppose	34	36	37	33	30
Reserve judgment	47	48	48	49	50
Don't know	1	0	1	1	1

	Cambridge 11/87 (%)	*Cambridge 2/89* (%)	*Gallup 2/89* (%)	*Cambridge 8/90* (%)
Favor	17	18	17	18
Oppose	34	32	23	27
Reserve judgment	47	48	59	54
Don't know	2	1	1	1

NOTE.—Sample sizes for Cambridge of approximately 1,500 until November 1989, and 1,250 thereafter. Sample sizes for Gallup of approximately 1,000.

Nuclear Pragmatism

11. Thinking about all energy sources available for large-scale use, would you say that nuclear energy is a good choice, a realistic choice, or a bad choice?

	Cambridge 5/86 (%)	*Cambridge 8/86* (%)	*Cambridge 8/86* (%)	*Cambridge 11/86* (%)	*Cambridge 2/87* (%)
Good choice	23	21	19	23	17
Realistic choice	52	57	48	52	50
Bad choice	23	21	31	23	32
Don't know	2	1	2	2	2

	Cambridge 5/87 (%)	*Cambridge 8/87* (%)	*Cambridge 8/87* (%)	*Gallup 9–10/87* (%)
Good choice	19	20	21	20
Realistic choice	51	53	49	47
Bad choice	28	25	29	29
Don't know	2	3	2	4

	Cambridge 11/87 (%)	*Cambridge 2/88* (%)	*Cambridge 8/88* (%)	*Cambridge 11/88* (%)	*Cambridge 2/89* (%)
Good choice	21	20	18	17	17
Realistic choice	47	49	49	50	50
Bad choice	27	27	29	30	29
Don't know	6	4	4	3	4

	Gallup *2/89* (%)	*Cambridge* *8/89* (%)	*Cambridge* *11/89* (%)	*Cambridge* *2/90* (%)	*Gallup* *2/90* (%)
Good choice	19	18	19	18	23
Realistic choice	50	49	48	44	45
Bad choice	25	29	31	32	26
Don't know	6	5	3	5	6

NOTE.—For four of the earlier Cambridge surveys (5/86, 8/86, 11/86, and 8/87), the second choice was worded "Not a good choice but a realistic choice." Sample sizes for Cambridge of approximately 1,500 until November 1989, and 1,250 thereafter. Sample sizes for Gallup of approximately 1,000.

12. I'm going to read you three statements about nuclear energy. Please tell me which of these statements most accurately reflects your own attitude toward nuclear energy:

		Cambridge	
	7/83 (%)	*11/83* (%)	*2/84* (%)
We should use the nuclear plants we now have, and build more nuclear plants	32	27	23
We should use the nuclear plants we now have, but not build any more plants	47	51	54
We should shut down all existing nuclear plants, and not build any more plants	19	21	20
Don't know	1	1	3
	5/84 (%)	*10/84* (%)	*1/85* (%)
We should use the nuclear plants we now have, and build more nuclear plants	31	30	29
We should use the nuclear plants we now have, but not build any more plants	47	51	46
We should shut down all existing nuclear plants, and not build any more plants	20	17	21
Don't know	2	3	3
	3/85 (%)	*5/85* (%)	*10/85* (%)
We should use the nuclear plants we now have, and build more nuclear plants	30	29	28
We should use the nuclear plants we now have, but not build any more plants	51	48	51
We should shut down all existing nuclear plants, and not build any more plants	17	19	17
Don't know	2	4	4
	2/86 (%)	*5/86* (%)	*8/86* (%)
We should use the nuclear plants we now have, and build more nuclear plants	29	26	28
We should use the nuclear plants we now have, but not build any more plants	52	52	51
We should shut down all existing nuclear plants, and not build any more plants	16	20	20
Don't know	3	3	1

	11/86	*2/87*	*5/87*
	(%)	(%)	(%)
We should use the nuclear plants we now have, and build more nuclear plants	26	26	26
We should use the nuclear plants we now have, but not build any more plants	51	51	52
We should shut down all existing nuclear plants, and not build any more plants	20	21	20
Don't know	2	2	2

	8/87	*11/87*	*5/87*
	(%)	(%)	(%)
We should use the nuclear plants we now have, and build more nuclear plants	28	24	30
We should use the nuclear plants we now have, but not build any more plants	49	51	49
We should shut down all existing nuclear plants, and not build any more plants	20	18	19
Don't know	3	6	2

	5/90	*11/90*	*2/91*
	(%)	(%)	(%)
We should use the nuclear plants we now have, and build more nuclear plants	24	33	32
We should use the nuclear plants we now have, but not build any more plants	53	43	48
We should shut down all existing nuclear plants, and not build any more plants	18	19	15
Don't know	4	5	5

Note.—Sample sizes of approximately 1,500 until November 1989, and 1,250 thereafter.

13. Many people have different opinions about building nuclear energy plants for providing for the future electricity needs of our country. Please tell me which statement comes closest to your opinion on the subject: We should build new nuclear energy plants now. We should not build new nuclear energy plants now but should keep the option open. We should not build new nuclear energy plants and should not keep the option open.

		B/G	
	5/91	*10/91*	*9/92*
	(%)	(%)	(%)
Build now	22	24	18
Keep option open	48	48	56
Should not build/keep option	24	22	23
Don't know	6	7	3

Note.—Sample sizes of approximately 1,000.

Nuclear Futures

14. Practically speaking, how important a role *should* nuclear energy play in meeting America's future energy needs? Should nuclear energy play a very important role, a somewhat important role, not too important a role, or not an important role at all? (Emphasis added.)

	Cambridge	*Cambridge*	*Cambridge*	*Gallup*	*Gallup*	*B/G*
	2/89	*5/89*	*2/90*	*7–8/91*	*2/92*	*3–93*
	(%)	(%)	(%)	(%)	(%)	(%)
Very important	36	38	35	39	35	32
Somewhat important	31	31	30	34	38	39
Not too important	13	13	14	11	10	12
Not important at all	15	14	15	14	12	12
Don't know	5	4	5	2	5	6

NOTE.—Sample sizes for Cambridge of approximately 1,500 until November 1989, and 1,250 thereafter. Sample sizes for Gallup and for B/G of approximately 1,000.

15. *How likely* do you think it is that new nuclear power plants will have to be built in the United States in the years ahead? Would you say it is very likely somewhat likely, not too likely, or not at all likely that the United States will have to build new nuclear power plants in the years ahead? (Emphasis added.)

	Cambridge	*Cambridge*	*Cambridge*[a]	*Cambridge*	*Cambridge*[a]
	5/88	*8/88*	*8/88*	*11/88*	*11/88*
	(%)	(%)	(%)	(%)	(%)
Very likely	40	42	36	43	39
Somewhat likely	38	35	35	35	32
Not too likely	12	12	18	13	17
Not at all likely	6	7	7	5	8
Don't know	3	4	4	5	4

	Cambridge	*Gallup*	*Cambridge*[a]	*Cambridge*	*Cambridge*[a]
	2/89	*2/89*	*2/89*	*5/89*	*5/89*
	(%)	(%)	(%)	(%)	(%)
Very likely	45	40	40	43	39
Somewhat likely	33	36	35	35	35
Not too likely	13	12	13	10	14
Not at all likely	4	6	8	7	8
Don't know	5	6	5	5	4

	Cambridge	*Cambridge*	*Cambridge*
	8/89	*11/89*	*2/90*
	(%)	(%)	(%)
Very likely	37	38	40
Somewhat likely	35	38	35
Not too likely	15	14	13
Not at all likely	8	8	8
Don't know	5	2	4

NOTE.—Sample sizes for Cambridge of approximately 1,500 until November 1989, and 1,250 thereafter. Sample size for Gallup of approximately 1,000.

[a] For four surveys a second version of the instrument was worded: "How likely do you think it is that new nuclear power plants will need to be built in the United States in the next 10 years? Would you say it is very likely, somewhat likely, not too likely, or not at all likely that the United States will need to build new nuclear power plants in the next 10 years?"

16. In order to meet the future power needs of the nation, how important do you feel it is to *have more* nuclear power plants—extremely important, somewhat important, not too important, or not at all important? (Emphasis added.)

		Gallup	
	7/76 (%)	*7/79* (%)	*6/86* (%)
Extremely important	34	30	14
Somewhat important	37	33	35
Not too important	10	13	21
Not at all important	8	17	26
No opinion	11	7	4

NOTE.—Sample sizes of approximately 1,000.

17. Do you think the nation's need for nuclear energy as part of the total energy mix will *increase* in the years ahead, or not? (Emphasis added.)

	Cambridge 11/86 (%)	*Cambridge 2/87* (%)	*Cambridge 5/87* (%)	*Gallup 8/87* (%)	*Cambridge 9–10/87* (%)
Yes	74	75	77	78	75
No	18	19	17	16	17
Not sure	8	6	6	7	8

	Cambridge 11/87 (%)	*Cambridge 2/88* (%)	*Cambridge 5/88* (%)	*Cambridge 8/88* (%)	*Cambridge 11/88* (%)
Yes	72	76	76	78	77
No	17	17	18	15	15
Not sure	11	7	6	7	8

	Cambridge 2/89 (%)	*Cambridge 5/89* (%)	*Cambridge 2/91* (%)	*Cambridge 11/91* (%)	*B/G 2/93* (%)
Yes	77	76	76	76	72
No	13	14	13	18	20
Not sure	9	10	11	7	8

NOTE.—Sample sizes for Cambridge of approximately 1,500 until November 1989, and 1,250 thereafter. Sample sizes for Gallup and for B/G of approximately 1,000.

18. How important do you think nuclear energy plants *will* be in meeting this nation's electricity needs in the years ahead? Do you think nuclear energy plants will be very important, somewhat important, not too important, or not important at all? (Emphasis added.)

	Cambridge 10/85 (%)	*Cambridge 2/86* (%)	*Cambridge 5/86* (%)
Very important	49	46	42
Somewhat important	28	28	29
Not too important	10	14	13
Not important at all	8	9	13
Don't know	4	4	3

	Cambridge 8/86 (%)	*Cambridge 11/86* (%)	*Cambridge 2/87* (%)
Very important	50	44	49
Somewhat important	28	25	30
Not too important	10	14	11
Not important at all	10	10	8
Don't know	2	6	2

	Cambridge 5/87 (%)	*Cambridge 8/87* (%)	*Gallup 9–10/87* (%)
Very important	53	52	47
Somewhat important	27	27	30
Not too important	10	9	13
Not important at all	8	8	6
Don't know	2	3	4

	Cambridge 11/87 (%)	*Cambridge 2/88* (%)	*Cambridge 5/88* (%)
Very important	49	53	55
Somewhat important	28	26	26
Not too important	11	10	10
Not important at all	8	7	7
Don't know	4	4	2

	Cambridge 8/88 (%)	*Cambridge 11/88* (%)	*Gallup 2/89* (%)
Very important	55	52	45
Somewhat important	25	27	34
Not too important	10	10	10
Not important at all	8	8	6
Don't know	3	3	5

	Cambridge 5/89 (%)	*Cambridge 8/89* (%)	*Cambridge 11/89* (%)
Very important	50	45	51
Somewhat important	27	29	29
Not too important	10	12	10
Not important at all	9	8	8
Don't know	4	6	3

	Cambridge 2/90 (%)	*Cambridge 5/90* (%)	*Cambridge 6/90* (%)
Very important	47	46	51
Somewhat important	32	31	32
Not too important	10	11	9
Not important at all	8	9	5
Don't know	4	4	3

	Gallup 7/90 (%)	*Cambridge 11/90* (%)	*Cambridge 2/91* (%)
Very important	44	47	44
Somewhat important	34	30	36
Not too important	8	10	9
Not important at all	9	7	7
Don't know	5	5	4

	B/G 2/93 (%)
Very important	45
Somewhat important	33
Not too important	8
Not important at all	7
Don't know	7

NOTE.—Sample sizes for Cambridge of approximately 1,500 until November 1989, and 1,250 thereafter. Sample sizes for Gallup and B/G of approximately 1,000.

Nuclear Compared with Other Energy Sources

19. In thinking about this country's future energy needs, which of these energy sources do you think we should rely on more for our future needs and which should we rely on less. . . . Nuclear power? Coal? Oil?

	Yankelovich					
	4/89			*4/91*		
	Nuclear Power (%)	*Coal* (%)	*Oil* (%)	*Nuclear Power* (%)	*Coal* (%)	*Oil* (%)
Rely more	45	42	33	44	43	27
Rely less	47	51	62	48	49	68
Not sure	8	7	5	8	8	5
N	1,012	1,012	1,012	1,012	1,012	1,012

20. Here is a list of possible energy sources. Please tell me which *one* source you think is. . . .

	Roper 1/85 (N = 1,989)				
	Nuclear (%)	*Natural Gas* (%)	*Coal* (%)	*Oil* (%)	*Don't Know* (%)
The safest to use?	6	46	26	16	7
Best for the environment?	13	62	8	8	10
The most economical for consumers?	14	44	24	8	11
The one we're least likely to run out of?	32	21	27	9	12
The one you find most acceptable for widespread use?	9	59	9	16	7

		Roper 1/85 (N = 1,989)			
	Nuclear (%)	*Natural Gas* (%)	*Coal* (%)	*Oil* (%)	*Don't Know* (%)
The one you find least acceptable for widespread use?	59	4	20	9	8
The one responsible for the most deaths through accidents over the past few years?	22	18	33	7	20
The one responsible for the fewest deaths through accidents over the past few years?	21	23	16	18	22
Potentially the most dangerous to human life?	82	5	4	1	8

		Roper 6/86 (N = 1,994)			
	Nuclear (%)	*Natural Gas* (%)	*Coal* (%)	*Oil* (%)	*Don't Know* (%)
The safest to use?	5	47	21	20	7
Best for the environment?	9	64	7	11	10
The most economical for consumers?	11	45	21	11	11
The one we're least likely to run out of?	29	23	24	10	15
The one you find most acceptable for widespread use?	7	58	8	18	9
The one you find least acceptable for widespread use?	59	4	20	7	9
The one responsible for the most deaths through accidents over the past few years?	43	13	16	6	21
The one responsible for the fewest deaths through accidents over the past few years?	15	23	22	16	24
Potentially the most dangerous to human life?	86	4	2	1	6

		Roper 1/90 (N = 1,994)			
	Nuclear (%)	*Natural Gas* (%)	*Coal* (%)	*Oil* (%)	*Don't Know* (%)
The safest to use?	7	61	13	10	9
Best for the environment?	13	64	5	5	12
The most economical for consumers?	12	54	14	6	14
The one we're least likely to run out of?	35	25	15	8	17
The one you find most acceptable for wide-spread use?	8	67	4	10	10
The one you find least acceptable for wide-spread use?	50	5	24	9	12
The one responsible for the most deaths through accidents over the past few years?	35	12	15	10	28
The one responsible for the fewest deaths through accidents over the past few years?	13	28	17	9	33
Potentially the most dangerous to human life?	76	4	5	2	12

		Roper 3/91 (N = 1,972)			
	Nuclear (%)	*Natural Gas* (%)	*Coal* (%)	*Oil* (%)	*Don't Know* (%)
The safest to use?	8	60	12	9	11
Best for the environment?	15	61	5	5	15
The most economical for consumers?	11	58	11	5	15
The one we're least likely to run out of?	28	29	15	9	19
The one you find most acceptable for wide-spread use?	10	64	6	9	11
The one you find least acceptable for wide-spread use?	48	5	26	7	13

	Roper 3/91 (N = 1,972)				
	Nuclear (%)	*Natural Gas* (%)	*Coal* (%)	*Oil* (%)	*Don't Know* (%)
The one responsible for the most deaths through accidents over the past few years?	34	9	14	10	33
The one responsible for the fewest deaths through accidents over the past few years?	12	31	13	9	35
Potentially the most dangerous to human life?	73	5	5	2	14

21. In generating electricity today, which one do you think is the cleanest and causes the least pollution . . . oil, natural gas, coal, or nuclear fuel?

	ORC			
	8/76 (%)	*7/80* (%)	*8/83* (%)	*12/89* (%)
Oil	18	23	6	5
Natural gas[a]	. . .	. . .	44	50
Coal	6	12	5	5
Nuclear	63	52	37	30
Don't know	13	13	8	10

NOTE.—Sample sizes of approximately 2,000.

[a] The natural gas option was not included in the 1976 and 1980 polls.

Danger and Risk

22. There are differences in opinion about how safe nuclear power plants are. Some people say they are completely safe, while others say they present dangers and hazards. How do you feel—that it would be safe to have a nuclear energy plant *someplace near here* or that it would present dangers? (Emphasis added.)

	Roper			
	10/73 (%)	*10/74* (%)	*10/75* (%)	*10/76* (%)
Safe	36	37	42	33
Would present dangers	41	41	43	46
Don't know	23	22	15	21

	9/77	*9/79*	*9–10/80*	*9/81*	*9/82*	*9/83*	*9/89*	*9/90*
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
Safe	39	27	28	26	21	26	22	25
Would present dangers	47	64	64	65	72	67	70	67
Don't know	14	9	8	9	7	7	8	8

NOTE.—Sample sizes of approximately 2,000.

23. We hear a lot these days about things that can be risks to people's health or safety. Here is a list of a number of them. Would you read over that list, and then for each one tell me whether it is something you think involves a high degree of risk to a person, or involves a moderate risk, or involves only a minor risk. . . . Living near a nuclear power plant.

	Roper							
	9/78	*4/79*	*6/79*	*4/81*	*9/85*	*5/86*	*9/88*	*9/92*
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
High risk	37	45	45	46	48	58	44	42
Moderate risk	31	32	27	27	28	25	30	34
Minor risk	26	19	24	22	21	15	22	21
Don't know	6	3	4	5	3	2	4	2
N	2,000	2,007	2,000	1,999	1,996	1,994	2,000	1,994

Nuclear as a Trade-off Choice

24. Now I'm going to read you two statements about the trade-off between nuclear energy versus foreign oil. Which of these views is closer to your opinion?

	Cambridge			
	1/85	*3/85*	*5/85*	*10/85*
	(%)	(%)	(%)	(%)
We should use more nuclear energy if it will reduce our dependence on foreign oil	55	60	60	58
We should use less nuclear energy even if it means increasing our dependence on foreign oil	32	29	30	29
Don't know	13	11	10	13

	2/86 (%)	5/86 (%)	8/86 (%)	11/86 (%)
We should use more nuclear energy if it will reduce our dependence on foreign oil	58	49	57	57
We should use less nuclear energy even if it means increasing our dependence on foreign oil	29	35	32	30
Don't know	14	16	11	13

	2/87 (%)	5/87 (%)	8/87 (%)	8/87 (%)
We should use more nuclear energy if it will reduce our dependence on foreign oil	57	57	58	58
We should use less nuclear energy even if it means increasing our dependence on foreign oil	32	31	30	28
Don't know	11	12	13	14

	8/87 (%)	11/87 (%)	11/87 (%)	11/87 (%)
We should use more nuclear energy if it will reduce our dependence on foreign oil	57[a]	58	56	59[a]
We should use less nuclear energy even if it means increasing our dependence on foreign oil	32[b]	27	28	27[b]
Don't know	12	14	15	13

	2/88 (%)	11/88 (%)	2/89 (%)	5/89 (%)	11/89 (%)
We should use more nuclear energy if it will reduce our dependence on foreign oil	60[a]	62[a]	55	57[a]	63
We should use less nuclear energy even if it means increasing our dependence on foreign oil	27[b]	26[b]	30	29[b]	28
Don't know	13	12	14	14	10

Note.—Sample sizes of approximately 1,500.

[a] Conditional clause worded, "We should use more nuclear energy if that will reduce our dependence on foreign oil" (rather than "if it will").

[b] Conditional clause worded, "We should use less nuclear energy even if that will increase our dependence on foreign oil" (rather than "if it means increasing").

25. Now I'm going to read you two statements about the trade-off between nuclear energy versus foreign oil. Which of these views is closer to your opinion? We should use more nuclear energy if that will reduce our dependence on foreign oil. We should use less nuclear energy even if that will increase our dependence on foreign oil.

	Gallup	
	9–10/87 (%)	*2/89* (%)
Use more	59	61
Use less	32	31
Don't know	9	8

NOTE.—Sample sizes of approximately 1,000.

26. *Gallup:* I'm going to read you some statements, and I'd like you to tell me, for each one, whether you agree or disagree with the statement.

a. Using nuclear energy to generate electricity cuts *greenhouse emissions and air pollution.*

	7/90 (%)	*7–8/91* (%)
Agree	52	63
Disagree	21	25
Don't know	27	12

b. We should use more nuclear energy if that will cut *greenhouse gas emissions and air pollution.*

	7/90 (%)	*7–8/91* (%)
Agree	74	71
Disagree	20	26
Don't know	6	3

NOTE.—Sample sizes of approximately 1,000.